

MATRICES AND THEIR APPLICATIONS

J. R. Branfield and H. W. Bell



MACMILLAN

Introductory
Monographs in Mathematics

MATRICES AND THEIR APPLICATIONS

INTRODUCTORY MONOGRAPHS IN MATHEMATICS

General Editor
A. J. MOAKES, M.A.

- | | |
|--|--|
| Numerical Mathematics
<i>Exercises in computing with a desk
calculating machine</i> | A. J. Moakes |
| Mathematics for Circuits | W. Chellingsworth |
| The Core of Mathematics
<i>An introduction to 'modern'
mathematics</i> | A. J. Moakes |
| A Boolean Algebra
<i>Abstract and Concrete</i> | A. P. Bowran |
| Programming by Case Studies
<i>An Algol Primer</i> | O. B. Chedzoy
and
Sandra E. Ford |

MATRICES AND THEIR APPLICATIONS

J. R. BRANFIELD and A. W. BELL

NOTTINGHAM COLLEGE OF EDUCATION

Macmillan Education

ISBN 978-1-349-81600-2 ISBN 978-1-349-81598-2 (eBook)
DOI 10.1007/978-1-349-81598-2

© J. R. Branfield and A. W. Bell 1970

Reprint of the original edition 1970

*All rights reserved. No part of this publication may
be reproduced or transmitted, in any form or by
any means, without permission*

First published 1970

Published by

MACMILLAN AND CO LTD

London and Basingstoke

*Associated companies in New York, Toronto,
Dublin, Melbourne, Johannesburg & Madras*

CONTENTS

	PAGE
<i>Preface</i>	vii
<i>Chapter</i> 1. Matrices and their Algebra	1
2. Transformations of the Plane	10
3. Further Transformation Techniques	27
4. Linear Equations	35
5. The Eigenvector	44
6. Application of the Eigenvector to Linear Mappings	55
7. Networks and Communications	68
8. Successive Events, Stochastic Processes and Markov Chains	89
9. Quadratic Forms	113
<i>Bibliography</i>	130
<i>Answers to Exercises</i>	131

PREFACE

THE power and adaptability of the matrix as a mathematical tool is evident by matrix techniques becoming used in an increasing number of fields. No mathematical education is now complete without some knowledge of their use.

The material in this book has come from a course in matrices and their applications given to students training as mathematics teachers. In each section, the starting point is an application or situation in which matrices arise. Techniques are then developed for dealing with the situation, and exercises and investigations are provided through which the reader may extend his understanding of the ideas.

We hope that students in sixth forms, Colleges of Education and of Technology, and in the first year at University will find this material useful. Teachers will find material suitable for extending their own knowledge of the subject, and also some situations which could form the basis of lessons for secondary school children.

We are indebted to our own students on whom we have tried out this material; to Dr. T. J. Fletcher (H.M.I.) for many of the ideas on which Chapters 7 and 8 are based; to the editor of the Monograph series, Mr. A. J. Moakes, for his generous help and encouragement and to the publishers.

Nottingham College of
Education, 1969.

J. R. BRANFIELD
A. W. BELL

1

MATRICES AND THEIR ALGEBRA

1.1 Introduction

Matrices were invented in the context of sets of linear equations, a topic which we shall discuss later, but the range of their applications has increased enormously in recent years, particularly in the planning of industrial operations, of production and marketing. We shall introduce matrices in relation to some of these newer applications.

1.2 The Matrix

There are many situations where it is natural to present data in the form of a table. One of these is shown in Table 1.1.

	English	Maths	French
Men	45	50	15
Women	55	20	20

Table 1.1. The matrix **F**. First-year students taking English, Mathematics or French at Nottingham College of Education.

A *matrix* is any rectangular array of numbers. The array of numbers obtained by omitting the headings in Table 1.1 is a matrix of two rows and three columns; this is called a ' 2×3 ' (two-by-three) matrix and will be denoted by **F**; we shall write it with round brackets thus:

$$\begin{pmatrix} 45 & 50 & 15 \\ 55 & 20 & 20 \end{pmatrix}.$$

A matrix consisting of only one row or column is called a *vector*; thus the column $\begin{pmatrix} 45 \\ 55 \end{pmatrix}$ would be called a column vector, and the row (45, 50, 15) would be called a row-vector. The individual numbers of a matrix are called its *entries* (or *components*).

1.3 Vector Addition

It is the object of matrix theory to develop methods of handling matrices as single units instead of having to deal with all their entries separately, in particular to give suitable meanings to addition and multiplication for matrices. This may be compared with what we do with fractions; the fraction $\frac{3}{4}$ consists of the pair of integers 3, 4; but we have developed ways of dealing with a fraction as a single unit, so that we know what we mean by the sum of two fractions, we know when to use this sum, and we know also how to obtain the sum from the two pairs of integers which constitute the original fractions: $(3, 4) + (1, 2) = (5, 4)$. Just as a fraction is thus a more complicated but more widely useful type of number than an integer, so a matrix is a more complicated but more widely useful type of number than either. We shall now consider what mathematical operations might be performed on the matrix $\mathbf{F}$, and hope to see whether any of these might usefully be regarded as addition or multiplication for matrices.

What further information can we gain from $\mathbf{F}$? Can we find the total number of students involved? Yes, this can be done by adding the numbers of men and women in each column, to give column totals; English 100, Mathematics 70, French 35, and grand total 205. This result may be checked by finding the row-totals first:

$$\begin{array}{ll} \text{men} & 45 + 50 + 15 = 110, \\ \text{women} & 55 + 20 + 20 = 95, \\ \text{grand total} & 110 + 95 = 205. \end{array}$$

These totals are included in the enlarged matrix $\mathbf{F}_2$ shown in Table 1.2.

	English	Maths	French	Row Totals
	e	a	f	r
Men	m 45	50	15	110
Women	w 55	20	20	95
Column Totals	c 100	70	35	205

Table 1.2. Table 1.1, with totals. The matrix $\mathbf{F}_2$.

Let us denote the row and column vectors of this matrix as follows:

Rows: Men **m**, Women **w**, row totals **r**

Columns: English **e**, Maths **a**, French **f**, column totals **c**

We shall define the sum of two vectors as the vector obtained by adding corresponding entries of the two vectors, and we can then write

what we have done as $\mathbf{m} + \mathbf{w} = \mathbf{c}$, this being a brief statement for $(45, 50, 15, 110) + (55, 20, 20, 95) = (100, 70, 35, 205)$ and $\mathbf{e} + \mathbf{a} + \mathbf{f} = \mathbf{r}$

which means
$$\begin{pmatrix} 45 \\ 55 \\ 100 \end{pmatrix} + \begin{pmatrix} 50 \\ 20 \\ 70 \end{pmatrix} + \begin{pmatrix} 15 \\ 20 \\ 35 \end{pmatrix} = \begin{pmatrix} 110 \\ 95 \\ 205 \end{pmatrix}.$$

Exercise

1.1 Compute the following vectors:

$$\mathbf{e} + \mathbf{a}, \quad \mathbf{e} - \mathbf{a}, \quad \mathbf{r} - \mathbf{e}, \quad \mathbf{m} + \mathbf{c}, \quad \mathbf{m} - \mathbf{w}.$$

The information given in Table 1.2 is lacking in one important respect – quite apart from the fact that it only quotes figures for three subjects. At this college, some students take two subjects, while others take only one. The row-totals do not therefore represent the total numbers of students involved in the three subjects, since those who take, for example, English *and* Mathematics will have been included twice, once under each heading. Table 1.3 shows the numbers of these, and omits the French students, for simplicity.

		English	Maths	both English and Maths	Row 'totals'
		$\mathbf{e}$	$\mathbf{a}$	$\mathbf{b}$	$\mathbf{r}_3$
Men	$\mathbf{m}_3$	45	50	2	93
Women	$\mathbf{w}_3$	55	20	3	72
Column Totals	$\mathbf{c}_3$	100	70	5	165

Table 1.3. The matrix $\mathbf{F}_3$.

Here the row totals are calculated so as to give the total number of students involved; the number taking both subjects must be subtracted from the sum of the separate English and Maths lists, since they will have been counted twice. In vector notation, we have

$$\mathbf{r}_3 = \mathbf{e} + \mathbf{a} - \mathbf{b}$$

using the letters indicated in the table.

1.4 Matrix addition

Table 1.4 shows the numbers of second-year students in the same categories as those of Table 1.1 (this was for first year students). The matrix obtained by omitting the headings here will be called $\mathbf{S}$,

		English	Maths	French
		<u>e</u>	<u>a</u>	<u>f</u>
Men	m	25	30	0
Women	w	35	18	0

Table 1.4. The Matrix **S**.

We shall define the sum $\mathbf{F} + \mathbf{S}$ of the two matrices $\mathbf{F}$ and $\mathbf{S}$ as the matrix formed by adding entries in corresponding positions, thus:

$$\begin{pmatrix} 45 & 50 & 15 \\ 55 & 20 & 20 \end{pmatrix} + \begin{pmatrix} 25 & 30 & 0 \\ 35 & 18 & 0 \end{pmatrix} = \begin{pmatrix} 70 & 80 & 15 \\ 90 & 38 & 20 \end{pmatrix}.$$

This gives us a matrix showing the total numbers in the first and second years combined, in each of the six categories.

1.5 Multiplication of a Matrix by a Scalar

Following a natural development we write $\mathbf{F} + \mathbf{F}$ as $2\mathbf{F}$, $\mathbf{F} + \mathbf{F} + \mathbf{F}$ to n terms as $n\mathbf{F}$; and we extend this definition to apply to multiplication by any number k drawn from the same set as the entries of the matrix thus:

The matrix $k\mathbf{A}$ is that obtained from $\mathbf{A}$ by multiplying each entry by k . Thus

$$3 \cdot 1 \begin{pmatrix} 0 & 0.2 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0.62 \\ -3.1 & 9.3 \end{pmatrix}.$$

This is called multiplying a matrix by a scalar; a single number is called a scalar to distinguish it from a vector and a matrix.

Exercises

Compute the following matrices:

1.2 $\mathbf{S}_3$ ($\mathbf{S}$ modified in the same way as $\mathbf{F}$ was to give $\mathbf{F}_3$); the $\mathbf{b}$ column is $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

1.3 $\mathbf{F} - \mathbf{S}$, $\mathbf{F}_3 - \mathbf{S}_3$, $4\mathbf{S}$.

1.6 Matrix multiplication

This refers to the multiplication of one matrix by another. We shall begin by investigating whether there is any way of combining two *vectors* which might reasonably be called multiplication.

Suppose the first-year students of Table 1.2 are to be entered for examinations. The examination fees per student are 40p for English, 30p for Mathematics (because it's easy to mark!) and 50p for French (to cover the extra cost of an oral). What will the total cost be? Using the bottom row of the table, it is, in pence,

$$100.40 + 70.30 + 35.50 = 4000 + 2100 + 1750 = 7850 \text{ (£78.50)}.$$

This is one possible way of multiplying the vectors (100, 70, 35), (the subject-numbers vectors) and (40, 30, 50), (the fee vector); it can be described as multiplying corresponding entries and adding the results. This product of two vectors is called the *inner product*, or sometimes the *scalar product*, since the result is a single number or scalar. The inner product of vectors **a** and **b** is denoted by **a** . **b**, which accounts for yet a third name, namely the *dot product* of the vectors **a** and **b**.

Exercises

- 1.4** If all students have to pay a games subscription, £3 per year for men and £2 per year for women, what total sum is paid by (a) English students, (b) Maths students, (c) all these students? (Use Table 1.3, **F**₃.) Show how these results may be obtained as inner products of the subscription vector **s** = (3, 2) with the appropriate vectors taken from Table **F**₃.
- 1.5** If **a** = (2, 1, 0), **b** = (-3, 0, 2), **c** = (0, 1, 0), form the inner products **a** . **b**, **b** . **c**, **a** . **c**.

Continuing the discussion of the last paragraph, we might wish to calculate the examination fees to be paid by men and women separately as well as the total. We should then be forming the inner product of the fee vector **v** = (40, 30, 50) with each of the rows of the matrix **F**₂ (Table 1.2), omitting the row totals, and recording the three results.

We have, for the men's fees,

$$\begin{aligned} (45, 50, 15) \cdot (40, 30, 50) &= 45.40 + 50.30 + 15.50 \\ &= 1800 + 1500 + 750 \\ &= 4050 \end{aligned}$$

and for the women's fees

$$\begin{aligned} (55, 20, 20) \cdot (40, 30, 50) &= 55.40 + 20.30 + 20.50 \\ &= 2200 + 600 + 1000 \\ &= 3800 \end{aligned}$$

and a total of 7850, as before. We record the results as a column vector

$\begin{pmatrix} 4050 \\ 3800 \\ 7850 \end{pmatrix}$, and the whole calculation would be written in matrix notation

$$\text{as } \begin{pmatrix} 45 & 50 & 15 \\ 55 & 20 & 20 \\ 100 & 70 & 35 \end{pmatrix} \begin{pmatrix} 40 \\ 30 \\ 50 \end{pmatrix} = \begin{pmatrix} 4050 \\ 3800 \\ 7850 \end{pmatrix}.$$

Here we may see good reason for writing the vector $\begin{pmatrix} 4050 \\ 3800 \\ 7850 \end{pmatrix}$ as a

column, since the components refer to men, women and total, and these are the titles of the rows of the matrix F_2 ; but it is not obvious why the fee-vector $(40, 30, 50)$, whose entries refer to the three subjects English, Maths, French should be written as a column rather than a row. The real answer to this cannot be given until we have more experience of matrix multiplication, but at least it can be seen that there must be some agreement about whether the vector $(40, 30, 50)$, when written following a matrix, is to be multiplied by the rows or the columns of the matrix. The agreement is that multiplication is always row by column, that is, each *row* of the *first* matrix is multiplied by each *column* of the *second*. (In this case, the 'second matrix' is a single column vector.) Also, in the product matrix, the row position of each product entry is that of the first matrix from which it was derived and its column position is that of the column of the second matrix from which it came. This may

be seen more clearly if we put another column alongside $\begin{pmatrix} 40 \\ 30 \\ 50 \end{pmatrix}$ to make a

3×2 matrix, and perform the multiplication:

$$\begin{pmatrix} 45 & 50 & 15 \\ 55 & 20 & 20 \\ 100 & 70 & 35 \end{pmatrix} \begin{pmatrix} 40 & 1 \\ 30 & -1 \\ 50 & 0 \end{pmatrix} = \begin{pmatrix} 4050 & -5 \\ 3800 & 35 \\ 7850 & 30 \end{pmatrix}.$$

Here the second *column* of the product matrix is obtained from the second *column* of the second matrix, and the *row* position of the entries in that product column is determined by the *row* of the first matrix from which they came.

In practice, we work like this:

$$\begin{array}{l} \begin{pmatrix} 45 & 50 & 15 \\ * & * & * \\ * & * & * \end{pmatrix} \begin{pmatrix} 40 & * \\ 30 & * \\ 50 & * \end{pmatrix} \begin{array}{l} 45.40 \\ \rightarrow + 50.30 \\ + 15.50 \end{array} \end{array} \left. \vphantom{\begin{pmatrix} 45 & 50 & 15 \\ * & * & * \\ * & * & * \end{pmatrix}} \right\} \begin{pmatrix} 4050 & * \\ * & * \\ * & * \end{pmatrix} \\ \text{then} \\ \begin{pmatrix} 45 & 50 & 15 \\ * & * & * \\ * & * & * \end{pmatrix} \begin{pmatrix} * & 1 \\ * & -1 \\ * & 0 \end{pmatrix} \begin{array}{l} 45.1 \\ \rightarrow + 50. -1 \\ + 15.0 \end{array} \end{array} \left. \vphantom{\begin{pmatrix} 45 & 50 & 15 \\ * & * & * \\ * & * & * \end{pmatrix}} \right\} \begin{pmatrix} 4050 & -5 \\ * & * \\ * & * \end{pmatrix} \\ \text{then} \\ \begin{pmatrix} * & * & * \\ 55 & 20 & 20 \\ * & * & * \end{pmatrix} \begin{pmatrix} 40 & * \\ 30 & * \\ 50 & * \end{pmatrix} \begin{array}{l} 55.40 \\ \rightarrow + 20.30 \\ + 20.50 \end{array} \end{array} \left. \vphantom{\begin{pmatrix} * & * & * \\ 55 & 20 & 20 \\ * & * & * \end{pmatrix}} \right\} \begin{pmatrix} 4050 & -5 \\ 3800 & * \\ * & * \end{pmatrix}$$

and so on.

We define A^2 as $A \cdot A$ and say that two matrices are *equal* if their corresponding entries are equal.

Exercises

1.6 If $\mathbf{A} = \begin{pmatrix} 3 & 5 \\ 1 & -1 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, compute the products of all possible pairs of these matrices, and $\mathbf{A}^2$, $\mathbf{B}^2$, $\mathbf{C}^2$.

1.7 If

$$\mathbf{E} = \begin{pmatrix} 2 & 3 & 5 \\ 1 & 1 & -1 \\ 1 & -2 & 1 \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} 0 & 2 & -1 \\ 3 & 3 & 3 \end{pmatrix},$$

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

compute the products of all possible pairs of these matrices, and $\mathbf{E}^2$, $\mathbf{F}^2$.

Note the effect of $\mathbf{C}$, $\mathbf{D}$, $\mathbf{G}$, $\mathbf{H}$ on the matrices they multiply.

1.7 Other ways of combining matrices

In this chapter so far we have taken a table of numbers and have discussed what mathematical operations might be performed in it, with a view to showing the operations of matrix addition and multiplication. We now look at another table with a view to seeing what mathematical operations are suggested – whether matrix addition and multiplication yield anything of interest, and whether any other operations emerge.

Table 1.5, taken from Volume II of the Crowther Report, shows the number of candidates obtaining scores 4, 3, 2, 1, 0 in an ability test, classified according to the number of children in the family in which they were brought up. The purpose of this table, in its context, was to demonstrate that the mean score showed a steady decrease as the size of family increased; the last row is the relevant one. In the calculation of this row a number of matrix operations can be observed. The first figure in the total score row is $4.5 + 3.14 + 2.9 + 1.24 + 0.13$; i.e. $\mathbf{c} \cdot \mathbf{c}_1$, and if the whole inner 5×6 matrix, omitting the headings and totals, is denoted by $\mathbf{M}$, the row $\mathbf{t}_s$ is the matrix product $\mathbf{cM}$. (Verify this.) The row $\mathbf{t}_c$ can also be written as a matrix product, if desired, as $\mathbf{uM}$, where $\mathbf{u}$ is the row vector $(1, 1, 1, 1, 1)$, and this might be useful in enabling both calculations to be dealt with by the same machine programme. The derivation of $\mathbf{m}$ from $\mathbf{t}_c$ and $\mathbf{t}_s$, by component-wise division, does not correspond to any matrix operation so far defined; we are obtaining a vector $(c_1, c_2, \dots, c_n)$ from $(a_1, a_2, \dots, a_n)$ and $(b_1, b_2, \dots, b_n)$ by putting $c_k = a_k/b_k$ for all k .

Expressing this in terms of multiplication, we have a product $b \vee c$ say (we invent the symbol $\vee$, calling it vee), defined for any two vectors of the same number of components by $(b \vee c)_k = b_k c_k$. This easily generalises to matrices of the same number of rows and columns and gives us a

c Score in ability test	c₁	c₂	c₃	c₄	c₅	c₆	t
	1	2	3	4	5	6 or more	Totals
s₄ 4	5	4	4	2	1	1	17
s₃ 3	14	24	32	23	16	22	131
s₂ 2	9	23	31	23	15	27	128
s₁ 1	24	31	34	39	37	42	207
s₀ 0	13	26	39	40	36	76	230
t_c Total cand.	65	108	140	127	105	168	713
t_s Total score	104	165	208	162	119	166	924
m Mean score	1.60	1.53	1.49	1.28	1.13	0.99	1.30

Table 1.5. Number of candidates obtaining different scores in an ability test, classified by family size.

(From the *Crowther Report*, Vol. II, p. 178.)

matrix product, $\mathbf{MvN}$ say, which corresponds to the matrix sum $\mathbf{M} + \mathbf{N}$ which we have already defined, multiplication simply replacing addition. The question of whether this new element-wise product $\mathbf{MvN}$ is more or less valid than the previous row-by-column product $\mathbf{MN}$ we leave over until we have completed our examination of the table above.

Exercises

- 1.8 Consider how each figure in the final totals column **t** is obtained, noting what matrix operations are used.
- 1.9 Make a similar study of Table 1.6, calculating totals and observing what operations are involved.

No. of children	Professional and managerial	Clerical	Skilled	Semi- skilled	Unskilled	Total
1	136	120	403	86	65	
2	277	201	772	172	108	
3	150	155	621	176	140	
4	78	77	442	108	127	
5	20	43	261	98	105	
6 or more	31	60	424	163	168	

Table 1.6. Number of families of different sizes in different socio-economic groups.

(From the *Crowther Report*, Vol. II.)

1.8 Laws of matrix Algebra

We now have one way of adding and two ways of multiplying matrices, so that we have made some progress towards our aim of enabling matrices to be manipulated like numbers. It is also clear from the exercises on p. 7 that the laws of matrix algebra differ from those of numbers in some respects – two matrices can only be combined if they have the right numbers of rows and columns, and even if matrices **A** and **B** *can* be multiplied (by matrix multiplication) in both orders, **AB** and **BA**, the product matrices are not in general the same.

At this point we must decide which of the two alternative multiplications to adopt, and we choose the row-by-column one as this is the one which is relevant to the applications we shall consider.

We can now summarise the laws of this matrix algebra.

First, it is convenient to define the multiplication of a matrix by a number: if λ is a number belonging to the same number system as that from which the entries of the matrix **M** are drawn, we define $\lambda\mathbf{M}$ to be the matrix obtained from **M** by multiplying every entry by λ .

The laws are as follows:

Associative Laws for Addition and Multiplication

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}; \quad \mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}.$$

Commutative Law (addition only)

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

Distributive Law, and laws concerning multiplication by a number:

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

$$\lambda(\mathbf{A} + \mathbf{B}) = \lambda\mathbf{A} + \lambda\mathbf{B}, \quad (\lambda + \mu)\mathbf{A} = \lambda\mathbf{A} + \mu\mathbf{A}$$

$$\lambda(\mathbf{AB}) = \lambda\mathbf{A} \cdot \mathbf{B}, \quad (\lambda\mu)\mathbf{A} = \lambda(\mu\mathbf{A}).$$

All of these can be verified from the definitions, most of them easily.

2

TRANSFORMATIONS OF THE PLANE

2.1 Use of 2×2 matrices

Matrices have an important part to play in the study of geometrical transformations. In this chapter we consider various transformations of the plane using 2×2 matrices; the investigation leads us to a number of important matrix ideas.

Consider a square piece of elastic material $OABC$ held at the origin of coordinates O (Fig. 2.1). It has a sleeve along OA and OC (through which the axes pass) so that when the material is pulled in the x and y axis directions, the material can move along the axes but will stay attached to them. The material is stretched by a factor of 2 in the x axis direction and by a factor of 3 in the y axis direction.

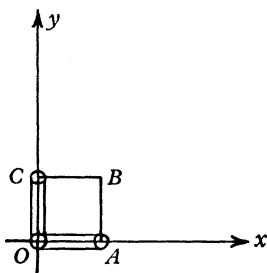


Fig. 2.1

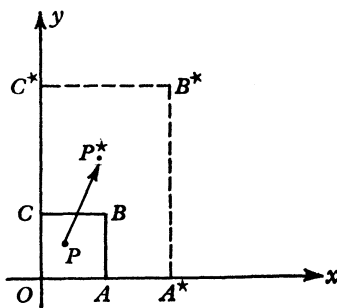


Fig. 2.2

By stretching, the square $OABC$ is transformed into the rectangle $OA^*B^*C^*$.

$$\begin{aligned} A(1, 0) &\rightarrow A^*(2, 0) \\ B(1, 1) &\rightarrow B^*(2, 3) \\ C(0, 1) &\rightarrow C^*(0, 3) \end{aligned}$$

and O remains fixed.

In general, a point $P(x, y)$ in the material will be transformed into the point $P^*(x^*, y^*)$ where

$$\left. \begin{aligned} x^* &= 2x \\ y^* &= 3y \end{aligned} \right\} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (1)$$

Equation (1) can be written as

$$x^* = 2x + 0y$$

$$y^* = 0x + 3y.$$

These may be put into matrix form as

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

– which may be verified by performing the matrix multiplication on the right-hand side.

Thus we see a close relationship between the matrix $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ and the transformation of the plane by stretches of factors 2 and 3 in the x and y directions respectively. In this chapter we shall be exploring this relationship, considering different transformations and the matrices to which they correspond. We shall be able to use matrix algebra in geometry whenever we have a transformation which can be considered as a set of linear equations. Such transformations of the plane as reflection in a line, rotation about a point, translation and so on can be treated as a set of linear equations and can thus be put into matrix form. Matrix algebra can be interpreted in geometrical terms whenever such interpretation is useful and helpful, and similarly geometrical transformations can be treated and manipulated by matrix algebra.

2.2 A matrix to rotate points through 90°

A point $P(x, y)$ is rotated through 90° anticlockwise about the origin O to the point $P^*(x^*, y^*)$. The point P is mapped onto the point P^* by a transformation of the plane which rotates points through 90° about the origin. From the geometry of the Fig. 2.3

$$x^* = -y$$

$$y^* = x.$$

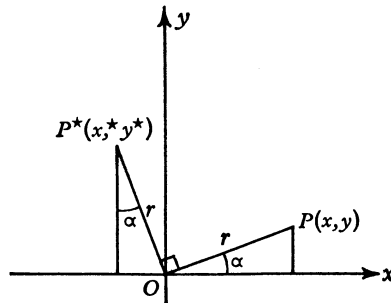


Fig. 2.3

The matrix form of these equations is

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Hence the matrix which will rotate points of the plane through 90° anticlockwise about the origin is

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

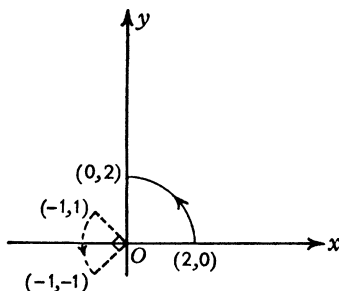


Fig. 2.4

For example the point $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ becomes $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ and the point $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ becomes $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$. The matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has the property of rotating *all* points of the plane through 90° about O , keeping the coordinate axes fixed, and can be considered as a transformation which rotates the *plane* through 90° .

2.3 A matrix to rotate the plane through any angle α

The plane is to be rotated through an angle α (anticlockwise) about the origin O , so that $P(x, y)$ goes to $P^*(x^*, y^*)$ with $OP = OP^*$ and $\angle POP^* = \alpha$. What is the matrix to perform this transformation?

To find a matrix which will perform a particular transformation we need to express the coordinates of the new position of a general point of the plane, in terms of its coordinates before the transformation; this requires a pair of equations.

Let the angle which OP makes with the x axis be θ ; and let OP and OP^* each be of length r (Fig. 2.5).

$$\begin{aligned} \text{Then } x^* &= r \cos(\alpha + \theta) \\ &= r \cos \alpha \cos \theta - r \sin \alpha \sin \theta; \end{aligned}$$

which since $x = r \cos \theta$ and $y = r \sin \theta$

becomes $x^* = x \cos \alpha - y \sin \alpha$.

Similarly $y^* = r \sin(\alpha + \theta)$

$$= r \sin \alpha \cos \theta + r \cos \alpha \sin \theta$$

and so $y^* = x \sin \alpha + y \cos \alpha$.

Hence the pair of equations is

$$x^* = x \cos \alpha - y \sin \alpha$$

$$y^* = x \sin \alpha + y \cos \alpha$$

from which we derive the matrix form

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Thus the geometric transformation represented by the matrix

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

maps the point $P(x, y)$ onto the point $P^*(x^*, y^*)$ by a rotation of the plane through an angle α about the origin. Putting $\alpha = 30^\circ$ gives the matrix to rotate the plane anticlockwise through an angle of 30° about O

$$\text{as } \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}.$$

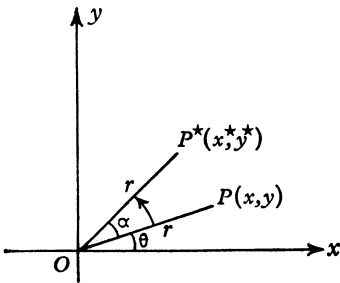


Fig. 2.5

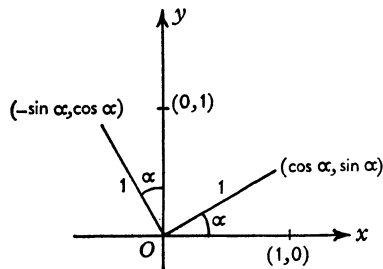


Fig. 2.6

Considering the transformation by this matrix of a rectangle with vertices $O(0, 0)$; $A(2, 0)$; $B(2, 1)$; $C(0, 1)$, we find the positions of the vertices of the transformed figure by using

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Putting in $(2, 0)$ for (x, y) we obtain the coordinates of A^* as

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix};$$

that is $x^* = \sqrt{3}$ and $y^* = 1$, so $A(2, 0) \rightarrow A^*(\sqrt{3}, 1)$. The reader should plot these coordinates and verify that the rectangle has been rotated anticlockwise through an angle 30° .

An alternative way of deriving the matrix for a rotation about the origin is as follows:

Consider the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

This maps $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ into $\begin{pmatrix} a \\ c \end{pmatrix}$, and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ into $\begin{pmatrix} b \\ d \end{pmatrix}$.

But under the rotation, (see Fig. 2.6)

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}.$$

The only possible matrix which can produce this transformation is therefore

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

Moreover, this matrix rotates *every* point of the plane about O in the same way, since

$$\begin{aligned} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \left[x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] \\ &= x \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

using the laws of matrix algebra (p. 9)

$$= x \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix} + y \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}$$

so that the right-angled triangle ONP has rotated into the position ON^*P^* , and thus OP has turned through the angle α about O . (Fig. 2.7).

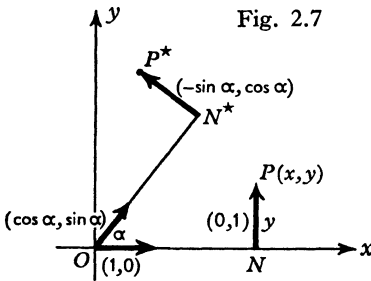


Fig. 2.7

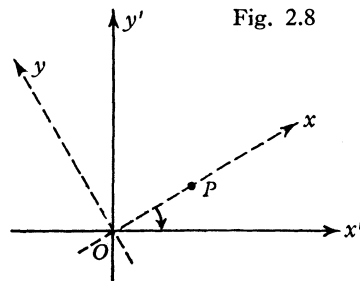


Fig. 2.8

2.4 Movement of figure or movement of axes ?

The equation

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

describes a transformation of (x, y) into (x^*, y^*) ; for example, if $(x, y) = (2, 0)$, then $(x^*, y^*) = (\sqrt{3}, 1)$. In the section above we have regarded (x, y) as denoting points of the rectangle in its initial position, and (x^*, y^*) , as the coordinates of the same points after the rectangle has been rotated. We might, however, adopt a different point of view and regard (x^*, y^*) , as the coordinates of the points of an unrotated rectangle with respect to *new axes*; axes which are obtained by rotating the old ones through 30° clockwise. The same matrix equation may represent either a rotation of the rectangle through $+30^\circ$, the axes remaining fixed, or a rotation of the axes through -30° , the rectangle remaining fixed. To put this another way, the coordinates (x, y) describe the position of points of a figure with respect to a given pair of axes. If (x, y) changes, then the position of the figure with respect to the axes changes; this may be regarded as a movement of the figure, or of the axes. In what follows a movement 'of the plane' means a movement of the figures in the plane, the axes remaining fixed. Thus the matrix for a rotation of the *axes* through a positive angle α is obtained from the equations on page 13 by changing α to $-\alpha$; this gives $\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$ as the matrix which keeps points of the plane fixed and refers them to new axes rotated through an angle α anticlockwise about O .

2.5 Other transformations in the plane

At the beginning of this chapter we considered the matrix $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ which may be described geometrically as a double-stretching of the plane, since for each point its distance from the y axis is multiplied by

2 and its distance from the x axis is multiplied by 3. This type of transformation will be called a 'double stretch'. The matrix $\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ has the geometric property of rotating the plane through an angle α about O and the matrix $\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$ describes the plane with respect to axes which are rotated through an angle α about O from an original set of axes. Other transformations of the plane to be considered are reflection, enlargement, shear and projection. The first two of these are straightforward and are dealt with in the exercises following.

Exercises

2.1 By finding suitable linear equations, find the matrix which will transform the plane so that each point of the plane

- (a) has its x -distance multiplied by a factor a ; y ordinate remains constant (linear stretch);
- (b) has its distance from the origin multiplied by a factor k (enlargement);
- (c) is reflected in the y axis (reflection).

2.2 Investigate the transformation of the rectangle $OABC$ with vertices $A(2, 0)$, $B(2, 1)$, $C(0, 1)$ when the plane is transformed by the matrices

$$\begin{array}{lll} \text{(a)} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, & \text{(b)} \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, & \text{(c)} \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}, \\ \text{(d)} \begin{pmatrix} 1 & -1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, & \text{(e)} \begin{pmatrix} 1 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -1 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \end{array}$$

2.6 The shear transformation

A shear may be considered as a transformation of the plane in which each point is moved along a line parallel to the x axis a distance proportional to its y -value. For the point $P(x, y)$, if a is the proportionality factor, then P is moved parallel to the x axis a distance ay . Thus $x^* = x + ay$, and since $y^* = y$, the matrix which describes this shear transformation is $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$.

Graph paper is needed in the following exercises in this chapter – it is important that the transformation of the plane by matrices is performed and seen.

Exercises

- 2.3 Using the shear matrix $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ show the transformation of the figures $OABC$, $ABDE$ and $BDFG$, where O is the origin, $A(2, 0)$; $B(2, 1)$; $C(0, 1)$; $D(-2, 1)$; $E(-2, 0)$; $F(-2, -1)$ and $G(2, -1)$.
- 2.4 Find the matrix which transforms the plane such that each point is moved along a line parallel to the y axis by a distance proportional to its x -value.
- 2.5 Investigate the transformations of the plane by the matrices (i) $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ and (ii) $\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix}$ – consider the transformation of the figures in 2.3.

2.7 Projection

A 'projection' will be considered as a transformation of the plane in which points are projected parallel to an axis onto a line through the origin.

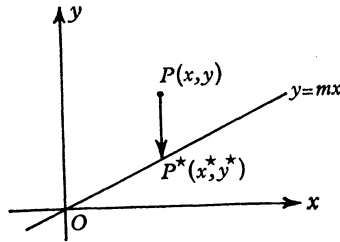


Fig. 2.9

For the point $P(x, y)$ its x coordinate remains unchanged in a projection parallel to the y axis. The y coordinate of P^* becomes mx for projection onto the line $y = mx$ and so

$$x^* = x$$

$$y^* = mx,$$

and the matrix which describes a projection of the plane onto the line

$$y = mx \text{ is } \begin{pmatrix} 1 & 0 \\ m & 0 \end{pmatrix}.$$

2.8 Unit and zero matrices

These two important matrices are connected with (i) the transformation which maps each point onto itself and (ii) the transformation which maps all points into the origin.

(i) If every point maps onto itself we have

$$x^* = x = 1 \cdot x + 0 \cdot y$$

and

$$y^* = y = 0 \cdot x + 1 \cdot y,$$

so that the matrix corresponding to this transformation is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

This is called the unit matrix and is denoted by **I**. In matrix algebra it plays the same part as the number 1 does in ordinary algebra. The mapping it performs is called the identity mapping.

(ii) If every point $P(x, y)$ is mapped into the origin $(0, 0)$, we have

$$x^* = 0 \cdot x + 0 \cdot y$$

and

$$y^* = 0 \cdot x + 0 \cdot y.$$

Thus the matrix which has the property of mapping all points into the origin is $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$; it is called the zero matrix.

2.9 Finding the transformation of a given matrix

Instead of finding a matrix to perform a particular transformation, we can also find what transformation a particular matrix performs on points of the plane.

Consider the matrix $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$; in particular, let us see how it transforms the rectangle $OABC$. (We could use a square, but this may be too

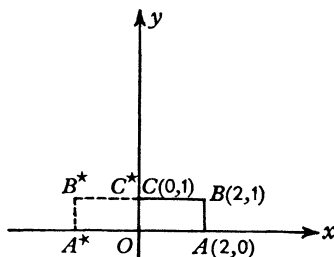


Fig. 2.10

regular to show the transformation – so, of course, may be the rectangle.) The equation for the transformation is $\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$, so that the point $B(2, 1)$ will be given by $\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$; that is $x^* = -2$ and $y^* = 1$, so that $B(2, 1)$ is mapped onto $B^*(-2, 1)$. Similarly $A(2, 0) \rightarrow A^*(-2, 0)$ and $C(0, 1) \rightarrow C^*(1, 0)$, Fig. 2.10. The general point $P(x, y)$ becomes $P^*(x, -y)$, so the matrix $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ has the geometric property of reflecting points of the plane in the axis of y .

2.10 Summary of plane transformations and corresponding 2×2 matrices

1. identity transformation $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (Unit matrix **I**)
2. zero transformation $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ (Zero matrix)
3. reflection (in x axis) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
4. rotation of figures through α about O $\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$
5. rotation (of axes) $\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$
6. enlargement (O fixed) $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$
7. linear stretch (parallel to x axis) $\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$
8. double stretch $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$
9. shear (parallel to x axis) $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$
10. projection (parallel to y axis, onto line $y = mx$), $\begin{pmatrix} 1 & 0 \\ m & 0 \end{pmatrix}$.

Exercises:

(Graph paper needed.)

- 2.6 (a) State the matrix which will enlarge a figure k times. Show that this matrix is k times the unit matrix.

- (b) Show diagrammatically what happens to the unit square with vertices $(0, 0); (1, 0); (1, 1); (0, 1)$, when $k=2, \frac{1}{2}, -3$.
- (c) Instead of the unit square find what transformation takes place, when $k=2$, when the figures are
- rectangle $(0, 0); (2, 0); (2, 1); (0, 1)$
 - circle $x^2 + y^2 = 1$
 - square $(-1, -1); (1, -1); (1, 1); (-1, 1)$.

2.7 For the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

show their effect on the rectangle in 2.6(c), (i) and identify the transformation of each of the matrices as one operation.

- 2.8 For the linear stretch matrix $\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$ let $a=2, \frac{1}{2}, -3$ and show the effect on the unit square and figures in 2.6(c), (i) and (ii).

In the instance of the circle, prove algebraically that the transformed circle is an ellipse.

- 2.9 Investigate the transformations described by the following matrices

$$(a) \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \quad (b) \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix}.$$

2.11 Transformations combined

In this section we consider whether ways of combining matrices – by addition, matrix multiplication and element-wise multiplication – give rise to any corresponding combination of the transformations which the matrices describe. It turns out that only one mode of combination is significant here – matrix multiplication.

Consider a point $P(x, y)$ rotated through 45° about O to $P^*(x^*, y^*)$ and then reflected in the axis of x to $P^{**}(x^{**}, y^{**})$, Fig. 2.11. From P to P^* we have (from the rotation of the plane matrix with $\alpha = 45^\circ$)

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad \dots \quad (1)$$

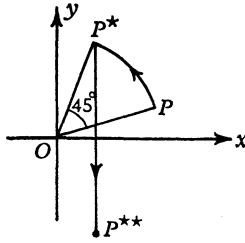


Fig. 2.11

From P^* to P^{**} we have

$$\begin{pmatrix} x^{**} \\ y^{**} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x^* \\ y^* \end{pmatrix}. \quad (2)$$

We seek now the single matrix which will describe the whole transformation.

Substituting from (1) for $\begin{pmatrix} x^* \\ y^* \end{pmatrix}$ into (2) then

$$\begin{aligned} \begin{pmatrix} x^{**} \\ y^{**} \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \left[\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right] \quad (3) \\ &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{(using the row on column method} \\ &\quad \text{for multiplying the matrices)} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned}$$

Thus the matrix which will perform a rotation through 45° followed by a reflection in the axis of x is

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Checking on the point $P(1, 1)$, Fig. 2.12, a rotation of 45° followed by a reflection in the axis gives P^{**} as $(0, -\sqrt{2})$. From the matrix we get

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -\sqrt{2} \end{pmatrix}.$$

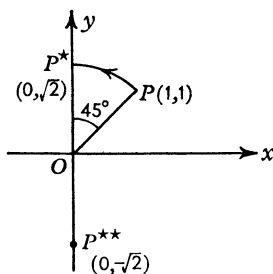


Fig. 2.12

The general proof that row on column matrix multiplication is needed is considered in the next section.

Exercise

2.10 Find the matrix corresponding to the performance of the two transformations above, in the reverse order.

2.12 Matrix multiplication and the combination of transformations

We see from the above exercise that the order of the matrices in (3) is important. Moreover, if **A** is the rotation matrix and **B** the reflection matrix, then from (3) the combined effect of performing **A** followed by **B** is given by **BA**; the matrix applied first being written second. (This has come about by the substitution used in placing $\begin{pmatrix} x^* \\ y^* \end{pmatrix}$ from (1) into (2).) For three transformations described by matrices **A**, **B** and **C** the combined effect of the matrices **A** followed by **B** followed by **C** will be given by computing **CBA**.

The process of matrix multiplication is directly related to following one transformation by another and is sometimes defined from this point of view. The working is as follows.

Let **A** be the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ which transforms the point $P(xy)$ into $P^*(x^*, y^*)$. The transformation can be written as a set of linear equations,

$$\begin{aligned} x^* &= ax + by \\ y^* &= cx + dy \end{aligned} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (4)$$

in matrix form $\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ or as $\mathbf{x}^* = \mathbf{Ax}$.

The transformation represented by **A** is followed by a transformation **B**, where **B** is the matrix $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$. By the matrix **B**, the point $P^*(x^*, y^*)$ is mapped onto $P^{**}(x^{**}, y^{**})$, and this can be written in matrix form as $\mathbf{x}^{**} = \mathbf{B}\mathbf{x}^*$.

From $\mathbf{x}^* = \mathbf{A}\mathbf{x}$ and $\mathbf{x}^{**} = \mathbf{B}\mathbf{x}^*$ then if we wish to obtain $\mathbf{x}^{**}$ in terms of the original coordinates $\mathbf{x}$, then substitution takes place and $\mathbf{x}^{**} = \mathbf{B}\mathbf{x}^* = \mathbf{B}(\mathbf{A}\mathbf{x}) = (\mathbf{BA})\mathbf{x}$. Note that the associative law for matrix multiplication is used here.

Now the transformation **B** can be put into a set of linear equations as

$$\left. \begin{aligned} x^{**} &= px^* + qy^* \\ y^{**} &= rx^* + sy^* \end{aligned} \right\} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (5)$$

Using equation (4) to find x^{**} and y^{**} in terms of x and y

$$\begin{aligned} x^{**} &= p(ax + by) + q(cx + dy) \\ y^{**} &= r(ax + by) + s(cx + dy), \end{aligned}$$

that is

$$\left. \begin{aligned} x^{**} &= (pa + qc)x + (pb + qd)y \\ y^{**} &= (ra + sc)x + (rb + sd)y \end{aligned} \right\} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (6)$$

The combined effect of the matrix **A** followed by the matrix **B** is given by equation (6). These in matrix form are:

$$\begin{pmatrix} x^{**} \\ y^{**} \end{pmatrix} = \begin{pmatrix} pa + qc & pb + qd \\ ra + sc & rb + sd \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

From this result and $\mathbf{x}^{**} = (\mathbf{BA})\mathbf{x}$

$$\begin{pmatrix} x^{**} \\ y^{**} \end{pmatrix} = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

we have

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} pa + qc & pb + qd \\ ra + sc & rb + sd \end{pmatrix}$$

giving the row on column method of matrix multiplication.

Exercises

(Graph paper needed.)

$$2.11 \quad \text{If} \quad \mathbf{M} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

evaluate the products **QM** and **MQ**. What transformations do these four matrices represent?

- 2.12** (a) With the matrix which will rotate points of the plane through 30° , show the transformation of the unit square. Transform this figure by a matrix which shears by a factor of 3 in the x axis direction. Find the matrix which will rotate points through 30° followed by a shear of factor 3 in the x axis direction. Use this matrix to transform the unit square. Show diagrammatically that the result of the separate transformations and of the combined transformation is the same.

(b) Find the matrix which will rotate points through 45° followed by a reflection in the x axis. Find the matrix which will reflect points in the x axis and then rotate them through 45° . Are the resultant matrices for these two sets of operations the same? Explain diagrammatically.

(c) Find the matrix which will enlarge the lengths of a figure to twice their size, then stretch it in the direction of the y axis by a factor of 3 and then reflect it in the line $x = y$.

- 2.13** Find the single transformation which is equivalent to reflection in $x = y$ followed by reflection in the x axis;

- (a) by sketching what happens to a figure.
(b) by multiplying the corresponding matrices.

- 2.14** Find the matrix which will rotate the plane first through an angle α and then through an angle β . The resultant matrix will be equivalent to the matrix which rotates the plane through an angle $\alpha + \beta$. Hence obtain the 'addition formulae':

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$$

$$\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta.$$

- 2.15** Find the matrix **A** which rotates points of the plane through 90° , and the matrix **B** which rotates points of the plane through 180° . Compute **AB** and **BA** and comment on your results algebraically and geometrically. Find the matrix **M** which reflects points of the plane in the line $y = x$ and the matrix **N** which reflects points of the plane in the y axis. Compute **MN** and **NM** and comment on your results algebraically and geometrically. Is there any relationship between the matrices **AB**, **MN**, **NM** – if so, why?

- 2.16** If $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$; $\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; $\mathbf{C} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$; investigate and comment on the transformations represented by **AB**, **BA**, **A(BC)** and **(AB)C**.

- 2.17 (a) Show the effect of the matrix $\begin{pmatrix} 2 & 4 \\ 0 & 2 \end{pmatrix}$ in transforming the unit square. Can the matrix be put in terms of the combined effect of basic matrices?

(b) Investigate the transformation effect of the matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}; \begin{pmatrix} 0 & -3 \\ 2 & 0 \end{pmatrix}; \begin{pmatrix} -1 & 0 \\ 3 & 1 \end{pmatrix}$$

(Note how the vertices correspond).

- 2.18 The reflections **A**, **B** take the point (x, y) into (y, x) and $(x, -y)$ respectively. Write down the matrix for each transformation.

Describe geometrically the effects of **AB**, **BA**, **AB**², **(BA)**², **ABA** and **BAB**.

Verify that

$$(\mathbf{AB})^4 = \mathbf{I} = (\mathbf{BA})^4 = (\mathbf{ABA})^2.$$

Write **ABABA** and **BABABA** using in each case four or less of the letters **A** and **B**.

- 2.19 A Digression – Combination tables for sets of transformations.

Consider the parallelogram shown, Fig. 2.13. The transformations which can be made which leave it occupying the same space in the plane are

- (i) a rotation (**R**) through 180° about *O*;
- (ii) no movement (the identity transformation).

These are known as the symmetries of the parallelogram.

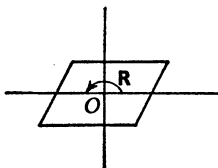


Fig. 2.13

Let us consider the four possible *combinations* of **R** and **I**, writing **IR** to mean **R** followed by **I** (p. 22 explains why this order is necessary). We display these in a table

		I	R	Transformation written second (performed first)
Transformation written first (performed second)	I	*		
	R			

in which the entry **IR** appears in the position marked *, that is in the **I** row and the **R** column.

Find the four entries in the table, in terms of the single transformation which would achieve the same final result as the two transformations performed successively. For example, **IR = R**, so the entry * is **R**.

Make a similar table for the addition of the integers 0, 1 modulo 2, and note that it has the same pattern as this one.

- 2.20** Give a list of the four symmetries of the rectangle, using the letters of Fig. 2.14: write the matrix corresponding to each symmetry. Form the table of combinations of these symmetries.

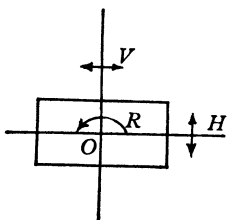


Fig. 2.14

If the reader is familiar with finite arithmetics, find among them a combination table which has the same pattern as this one.

3

FURTHER TRANSFORMATION TECHNIQUES

3.1 The inverse matrix

The matrix which gives an enlargement of a figure by a factor of k is $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} = \mathbf{A}$ say.

Is there a matrix which will restore the figure to its original state? Such a matrix will be $\begin{pmatrix} 1/k & 0 \\ 0 & 1/k \end{pmatrix} = \mathbf{B}$ which reduces a figure by $1/k$ times. The effect of $\mathbf{A}$ followed by $\mathbf{B}$ is to leave the figure unchanged.

We may also verify that the product of these matrices is $\mathbf{I}$

$$\begin{pmatrix} 1/k & 0 \\ 0 & 1/k \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A matrix $\mathbf{M}$ is called the *inverse* of the matrix $\mathbf{A}$, if $\mathbf{AM} = \mathbf{MA} = \mathbf{I}$. Thus $\mathbf{B}$ is the inverse of $\mathbf{A}$.

3.2 Inverses of some known matrices

What is the inverse of the matrix $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$?

The matrix $\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ reflects points in the y axis.

The matrix which will restore points to their original position will need to reflect in the y axis again; thus the inverse of $\mathbf{A}$ is $\mathbf{A}$ itself and

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We can show that if $\mathbf{M}$ is a left-inverse of $\mathbf{A}$, i.e. $\mathbf{MA} = \mathbf{I}$, then $\mathbf{M}$ is also a right inverse, i.e. $\mathbf{AM} = \mathbf{I}$. Let $\mathbf{Y}$ be a right inverse so $\mathbf{AY} = \mathbf{I}$.

In $\mathbf{AY} = \mathbf{I}$ premultiply each side by $\mathbf{M}$; then $\mathbf{MAY} = \mathbf{MI}$ and hence $\mathbf{MAY} = \mathbf{M}$.

But since $\mathbf{MA} = \mathbf{I}$, and $\mathbf{MAY} = (\mathbf{MA})\mathbf{Y}$ by the associative law, $\mathbf{IY} = \mathbf{M}$ thus $\mathbf{Y} = \mathbf{M}$ and the left and right inverses are the same.

The matrix and its inverse thus commute. This can be seen geometrically for if $\mathbf{A}$ followed by $\mathbf{M}$ restores points to their original position, so $\mathbf{M}$ followed by $\mathbf{A}$ will keep points where they are; the two operations are inverse. $\mathbf{M}$ can be written as $\mathbf{A}^{-1}$, i.e. $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$.

3.3 Inverse of the general 2×2 matrix

What is the inverse of the matrix $\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$?

Here we cannot deduce the inverse from geometrical considerations.

Let the inverse $\mathbf{M}$ of $\mathbf{A}$ be $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$.

Then as $\begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ gives

$$\left. \begin{array}{l} 3p + 5q = 1 \\ p + 2q = 0 \\ 3r + 5s = 0 \\ r + 2s = 1 \end{array} \right\} \text{ we obtain } \begin{array}{l} p = 2 \\ q = -1 \\ r = -5 \\ s = 3 \end{array}$$

by solving the equations.

$\mathbf{M}$ is thus $\begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}$

$$\left[\text{check that } \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right.$$

$$\text{and also } \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \left. \right].$$

Exercise

3.1 Find the inverse matrix of

$$(i) \begin{pmatrix} 4 & 1 \\ 7 & 2 \end{pmatrix}; \quad (ii) \begin{pmatrix} 3 & 1 \\ -7 & -2 \end{pmatrix}; \quad (iii) \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}.$$

Is any pattern becoming noticeable?

To find the inverse of the general 2×2 matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Let the inverse of $\mathbf{A}$ be $\mathbf{M} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}$ and it is required to find p, q, r, s in terms of a, b, c, d .

From $\begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ we obtain

$$\begin{aligned}ap + qc &= 1 \\bp + qd &= 0 \\ra + sc &= 0 \\rb + sd &= 1.\end{aligned}$$

By solving these equations $p = d/(ad - bc)$, $q = -b/(ad - bc)$, $r = -c/(ad - bc)$ and $s = a/(ad - bc)$. The inverse matrix $\mathbf{A}$ is thus $1/(ad - bc) \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Note the pattern of the elements.

The expression $(ad - bc)$ is known as the *determinant* of the matrix. $\mathbf{A}$ is $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and its determinant is denoted by $\det \mathbf{A}$ or by $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Its value $(ad - bc)$ is found by multiplying the entries in the principal or leading diagonal and subtracting the product of the entries in the other diagonal. Note that in the initial examples the determinant value was 1.

Exercise

3.2 Write down the values of the determinants of the following matrices, then write down their inverses and check your answers

- (i) $\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$; (ii) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$; (iii) $\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$; (iv) $\begin{pmatrix} 3 & 2 \\ 1 & 5 \end{pmatrix}$;
(v) $\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = \mathbf{R}$. Show the inverse of $\mathbf{R}(\mathbf{R}^{-1})$ is such that the rows of $\mathbf{R}$ are the columns of $\mathbf{R}^{-1}$. $\mathbf{R}$ said to be *transposed*. To transpose a matrix $\mathbf{A}$, rows of $\mathbf{A}$ become columns of the transposed matrix, denoted by $\mathbf{A}'$. Then $\mathbf{R}^{-1} = \mathbf{R}'$.

3.4 The uniqueness of the inverse

The matrix inverse to $\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$ is $\frac{1}{2} \begin{pmatrix} 3 & -1 \\ -4 & 2 \end{pmatrix}$. Is there another matrix which is inverse to $\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$?

Let $\mathbf{A}$ be the matrix and suppose $\mathbf{X}$ and $\mathbf{Y}$ are inverses of $\mathbf{A}$, then $\mathbf{XA} = \mathbf{AX} = \mathbf{I}$ and $\mathbf{YA} = \mathbf{AY} = \mathbf{I}$.

From $\mathbf{AX} = \mathbf{I}$

$$\begin{aligned}\mathbf{Y}(\mathbf{AX}) &= \mathbf{Y}(\mathbf{I}) && \text{premultiplying each side by } \mathbf{Y} \\(\mathbf{YA})\mathbf{X} &= \mathbf{Y} && \text{associative law for matrix multiplication} \\ \mathbf{IX} &= \mathbf{Y} && \text{since } \mathbf{YA} = \mathbf{I} \\ \mathbf{X} &= \mathbf{Y}.\end{aligned}$$

The inverse of a matrix is unique.

3.5 Matrices with no inverse

What is the inverse of $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$? As we follow the process for finding an inverse as in the previous section, the inverse of the matrix $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ will have a multiplying factor of the reciprocal of the determinant value of the matrix, $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. This determinant value is $1 \cdot 4 - 2 \cdot 2 = 0$ and so as the reciprocal of 0 is not determined, the matrix $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ has no inverse. All matrices of the form $\begin{pmatrix} a & b \\ ka & kb \end{pmatrix}$ will have zero determinants and thus no inverse. The important geometrical significance of a matrix which has or does not have an inverse is discussed in the next section. A matrix which has no inverse is said to be *singular*, whilst the matrix which has an inverse (i.e. its determinant is not zero) is said to be *non-singular*.

Exercise

3.3 Show diagrammatically the figure into which the unit square is transformed by the matrix $\mathbf{A} = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$. Similarly show the transformation of the unit square by the matrix $\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. What is the geometrical difference between the transformations of $\mathbf{A}$ and $\mathbf{B}$?

In the case of $\mathbf{B}$, show all points of the plane are mapped onto the line $y = 2x$. The point $(1, 1)$ goes to $(3, 6)$. What other points are also mapped onto $(3, 6)$?

3.6 Relations between determinants of matrices

Exercise

3.4 To investigate whether the determinant of two matrices $\mathbf{A}$ and $\mathbf{B}$ are related in any way.

Choose any two 2×2 matrices $\mathbf{A}$ and $\mathbf{B}$.

- (i) Find $\det \mathbf{A}$ and $\det \mathbf{B}$; then find $\mathbf{AB}$ and $\det \mathbf{AB}$. Show $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$.
- (ii) Similarly find $\mathbf{A} + \mathbf{B}$ and $\det (\mathbf{A} + \mathbf{B})$. Show that in general $\det (\mathbf{A} + \mathbf{B}) \neq \det \mathbf{A} + \det \mathbf{B}$.
- (iii) Compare $\det \mathbf{A}$ with $\det 3\mathbf{A}$ and show $\det 3\mathbf{A} = 3^2 \det \mathbf{A}$.

3.7 The translation matrix

We have so far concerned ourselves with such basic geometric transformations of the plane as reflection, rotation, enlargement and shear, and the combination of these. A notable transformation omitted in our list is translation. All 2×2 matrices preserve the origin, but in a translation the origin will be shifted.

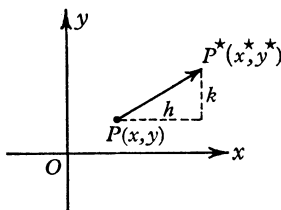


Fig. 3.1

The plane is to be translated so that the x coordinates are increased by a constant h , and the y coordinates by k . The general point $P(x, y)$ in the plane is mapped onto $P^*(x^*, y^*)$ where

$$\begin{aligned} x^* &= x + h \\ y^* &= y + k \end{aligned} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (1).$$

We have now to extract the translation matrix from these equations – which are linear but have a constant at the end; a situation we have not met before.

Equation (1) can be written in matrix form as

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} 1 & 0 & h \\ 0 & 1 & k \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (2)$$

giving as our translation matrix $\mathbf{T} = \begin{pmatrix} 1 & 0 & h \\ 0 & 1 & k \end{pmatrix}$. This is a matrix 2×3 matrix which will be awkward to fit and interpret with our 2×2 matrices. Matrices which are square and of the same order are easier to handle. We cannot reduce our $\mathbf{T}$ matrix to a 2×2 , but can increase it to a 3×3 as

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix}.$$

The form (2) then has to become $\begin{pmatrix} x^* \\ y^* \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$ and the

equations are

$$x^* = x + h$$

$$y^* = y + k$$

$$1 = 1$$

where the last equation is there to see things right!

Since $\mathbf{T}$ is a 3×3 matrix, its combination with 2×2 matrices can only be achieved if we increase the order of our 2×2 to 3×3 . We do this by use of the 1 in the bottom right hand corner: e.g. the rotation matrix

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \text{ becomes } \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix which translates the axes so that points of the plane remain fixed and the new origin is (r, s) is $\begin{pmatrix} 1 & 0 & -r \\ 0 & 1 & -s \\ 0 & 0 & 1 \end{pmatrix}$.

Exercises

- 3.5** Find the matrix which will rotate the plane about the origin through an angle α , followed by a translation of the plane through values (h, k) . Give some value to α , h and k and check geometrically that the result is correct.
- 3.6** $\mathbf{T}_x$ is a translation of 2 units in the x axis direction and $\mathbf{T}_y$ is a translation of 3 units in the y axis direction. $\mathbf{M}_x$ and $\mathbf{M}_y$ are reflections on the x and y axes respectively; and $\mathbf{R}$ is a rotation through 90° anticlockwise about the origin. P is the point (a, b) . Find the coordinates of P_1, P_2, P_3, P_4 formed from the point P under the following transformations
- (i) P_1 from P by $\mathbf{T}_x$ followed by $\mathbf{M}_x$.
 - (ii) P_2 from P by $\mathbf{T}_y$ followed by $\mathbf{M}_x$.
 - (iii) P_3 from P by $\mathbf{R}$ followed by $\mathbf{T}_x$.
 - (iv) P_4 from P by $\mathbf{R}$ followed by $\mathbf{M}_y$.
- 3.7** Find the matrix which will reflect the plane in the line $y = mx + c$. (Hint: translate the axes to $(0, c)$ followed by the matrix to reflect the plane in the line $y = mx$, followed by translation of the axes to $(0, -c)$.)

3.8 3×3 matrices as transformations of space

As yet we have considered 2×2 matrices and in the main we shall concern ourselves with matrices of this small size. The only reason for doing this is simplicity! It is easy to work with 2×2 matrices and the methods and ideas used can, in general, be extended to higher orders of matrices. We have with 2×2 matrices, transformation of the plane and it is straightforward to extend these methods to 3×3 matrices as transformation of space: we may even extend these to 4×4 matrices as transformation in 4 dimensions.

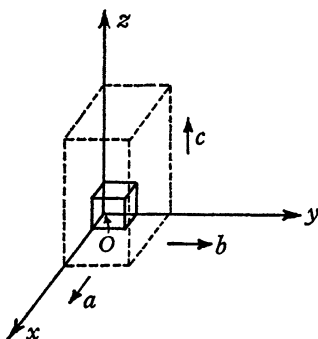


Fig. 3.2

In 3 dimensions we have the three mutually perpendicular axes x , y and z . If we stretch space by a factor of a in the x axis direction, a factor of b in the y axis direction and a factor of c in the z axis direction, then a point $P(x, y, z)$ is mapped onto the point $P^*(x^*, y^*, z^*)$ where

$$x^* = ax$$

$$y^* = by$$

$$z^* = cz.$$

The stretch matrix is thus

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}.$$

Our cube in Fig. 3.2 becomes a cuboid.

Exercises

3.8 Find the 3×3 matrices which transform space such that

- (i) reflection takes place in the $x - y$ axes plane
- (ii) enlargement by a factor of k

- (iii) a shear is performed with a factor of a in the x axis direction and a factor of b in the y axis direction (a diagram to see the result of the transformation will help before proceeding to the equations).

3.9 Describe geometrically the transformation of space by the following matrices:

$$\begin{array}{lll}
 \text{(i)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \text{(ii)} \begin{pmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \text{(iii)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \text{(iv)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} & &
 \end{array}$$

3.9 Inverses of matrices of higher order than 2×2

Again we have only considered inverses of 2×2 matrices. It is possible to find the inverse of a 3×3 matrix by solving sets of equations using the same idea as used in Page 28. It is instructive to do this (although the equations are heavier) and the 3×3 determinant of a matrix is obtained. However, an easier method of finding inverses of higher order matrices will be given in the next chapter.

4

LINEAR EQUATIONS

4.1 Simultaneous linear equations

The solution of two simultaneous linear equations

$$3x + 2y = 8$$

$$4x + 3y = 11$$

is as follows.

These equations can be written as $\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 8 \\ 11 \end{pmatrix}$, that is if $\mathbf{A}$ is the matrix $\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}$ and $\mathbf{k}$ is the vector $\begin{pmatrix} 8 \\ 11 \end{pmatrix}$, then the set of linear equations can be written as $\mathbf{Ax} = \mathbf{k}$.

Premultiply each side by $\mathbf{A}^{-1}$; then

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{k},$$

$$\mathbf{Ix} = \mathbf{A}^{-1}\mathbf{k},$$

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{k}.$$

To solve the given set of equations we need to premultiply $\begin{pmatrix} 8 \\ 11 \end{pmatrix}$ by the inverse of $\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}$.

The inverse of $\mathbf{A}$ is $\begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix}$, so

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix} \begin{pmatrix} 8 \\ 11 \end{pmatrix}.$$

Thus

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \text{or} \quad \begin{cases} x=2 \\ y=1 \end{cases}.$$

There are, of course, more direct methods of solving these equations. We give the matrix method here as an example of a method which generalises immediately to any number of equations.

Exercises

4.1 By matrix methods solve these sets of equations, and check your answers by substitution:

$$\begin{array}{lll} \text{(a)} & 4x + 5y = 23 & \text{(b)} \quad 3x - y = 11 \quad \text{(c)} \quad 2x + 3y = 4 \\ & x + 2y = 8 & 2x + 5y = -4 \quad -6x + 3y = 8 \end{array}$$

4.2 By matrix methods solve the sets of equations given by $\mathbf{Ax} = \mathbf{h}$

where $\mathbf{A} = \begin{pmatrix} 6 & 2 \\ 11 & 4 \end{pmatrix}$ and $\mathbf{h}$ takes in succession the values

$$\text{(i)} \begin{pmatrix} 14 \\ 26 \end{pmatrix} \quad \text{(ii)} \begin{pmatrix} 10 \\ 18 \end{pmatrix} \quad \text{(iii)} \begin{pmatrix} 7 \\ 13 \end{pmatrix} \quad \text{(iv)} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{(v)} \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

(vi) own choice of values.

4.2 The nature of solutions

As we have shown, the solution to the set of equations

$$3x + 2y = 8$$

$$4x + 3y = 1$$

is $x = 2$, and $y = 1$.

What would be the solution if the equations were homogeneous; that is

$$3x + 2y = 0$$

$$4x + 3y = 0?$$

The only solution is $x = y = 0$.

The matrix $\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$ is singular (since the determinant is zero) so this matrix has no inverse. What is the solution to the set of equations

$$2x + 4y = 6$$

$$x + 2y = 3?$$

The solutions are $x = 3 - 2t$, $y = t$ for any real number t ; an infinite set of solutions.

What is the solution if these equations are homogeneous, i.e. $\mathbf{h} = \mathbf{0}$? This time we have solutions $x = 2t$, $y = -t$, again an infinite set of solutions.

The equations

$$2x + 4y = 6$$

$$x + 2y = 2,$$

have no solution,

The nature of the solutions of a set of linear equations can be summarised as

	non-homogeneous ($\mathbf{h} \neq \mathbf{0}$)	homogeneous ($\mathbf{h} = \mathbf{0}$)
(i) non singular matrix ($\det \mathbf{A} \neq 0$)	unique solution	solutions = 0
(ii) singular matrix ($\det \mathbf{A} = 0$)	many solutions or no solutions (depending on $\mathbf{h}$ values)	many solutions

Exercise

4.3 Make sets of equations in each of the above five cases and show the graphical interpretation for solutions.

The part of this work which will be required later is that in a set of homogeneous equations, if solutions *other than zero* are required the equations must be such that $\det \mathbf{A} = 0$, i.e. the matrix must be singular.

4.3 Inverse of a 3×3 matrix

The set of simultaneous linear equations

$$\left. \begin{array}{l} x - y + 2z = 6 \\ 3x - 2y + z = 7 \\ x + 3y - z = -4 \end{array} \right\} \quad . \quad . \quad . \quad . \quad (1)$$

is solved conventionally by multiplying the first equation by -3 and adding the result to the second equation; multiplying the first equation by -1 and adding to the last equation to give

$$\left. \begin{array}{l} x - y + 2z = 6 \\ y - 5z = -11 \\ 4y - 3z = -10 \end{array} \right\} \quad . \quad . \quad . \quad . \quad (2)$$

The set of equations (2) is equivalent to the set (1).

We continue to solve the equations by multiplying the second equation by -4 and adding the result to the last equation, to give

$$\left. \begin{array}{l} x - y + 2z = 6 \\ y - 5z = -11 \\ 17z = 34 \end{array} \right\} \quad . \quad . \quad . \quad . \quad (3)$$

From here we obtain the value of z and, by substitution, the values of first y and then x as

$$\left. \begin{array}{l} x=1 \\ y=-1 \\ z=2 \end{array} \right\} \quad . \quad . \quad . \quad . \quad (4).$$

This method of solution, reducing the set of equations to a triangular form and ultimately to a diagonal form (4) gives an indication of the process to be used with matrices. (It also provides a systematic method for solving sets of linear equations.) The set of equations (1) in matrix form is

$$\begin{pmatrix} 1 & -1 & 2 \\ 3 & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 7 \\ -4 \end{pmatrix} \quad \text{or} \quad \mathbf{Ax} = \mathbf{h}$$

and the solution set (equations (4)) is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

The process of solution can be regarded as the method by which

$$\begin{pmatrix} 1 & -1 & 2 \\ 3 & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix} \text{ is changed into } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

by a process called row reduction – a sequence of operations performed on the rows of the matrix.

Consider the matrix $\mathbf{E}_1 = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$ with the matrix

$\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ 3 & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix}$, noting that $\mathbf{E}_1$ is basically the 3×3 identity matrix

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ with the second two elements of the first column of $\mathbf{A}$ used

with changed sign.

Then $\mathbf{E}_1\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 3 & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 4 & -3 \end{pmatrix}$ gives the

left-hand set of equations (2). The matrix $\mathbf{E}_1$ has the property of performing the row reduction required in solving the equations since the row $(-3 \ 1 \ 0)$ in $\mathbf{E}_1$, when matrix multiplied with $\mathbf{A}$ multiplies row 1

of $\mathbf{A}$ by -3 and adds it to row 2 of $\mathbf{A}$ – the method used at the beginning. By using $\mathbf{E}_1$ with $\mathbf{h}$ as $\mathbf{E}_1\mathbf{h}$ the right-hand side of equations (2) is obtained.

Continuing to solve we would choose $\mathbf{E}_2$ as $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix}$ and use it

before $\mathbf{E}_1\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 4 & -3 \end{pmatrix}$ from above; so

$$\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 4 & -3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 17 \end{pmatrix}.$$

Thus $\mathbf{E}_2\mathbf{E}_1\mathbf{Ax} = \mathbf{E}_2\mathbf{E}_1\mathbf{h}$ and gives the set of equations (3). At this stage we solved the equations by finding z and substituting. We could continue using matrices.

Choose $\mathbf{E}_3$ as $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{17} \end{pmatrix}$ then

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{17} \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 17 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{pmatrix} = (\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}).$$

We now chose $\mathbf{E}_4$ as $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}$ (note the basic identity matrix with the 5). Then

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}$$

and the right-hand side would be $\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{h}$ and would give the value of y .

Our final choice for $\mathbf{E}_5$ to obtain x would be $\begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then

$$\begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}.$$

Thus by a series of matrices $\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1$ we have changed the original matrix $\begin{pmatrix} 1 & -1 & 2 \\ 3 & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix}$ into $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ i.e. $\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1$ must be the *inverse* of $\mathbf{A}$ (as $\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{I}$).

$$\begin{aligned}
 \mathbf{E}_5 \mathbf{E}_4 \mathbf{E}_3 \mathbf{E}_2 \mathbf{E}_1 &= \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{17} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix} \\
 &\quad \times \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} -\frac{1}{17} & \frac{5}{17} & \frac{3}{17} \\ \frac{4}{17} & -\frac{3}{17} & \frac{5}{17} \\ \frac{11}{17} & -\frac{4}{17} & \frac{1}{17} \end{pmatrix} = \mathbf{E} \text{ say.}
 \end{aligned}$$

Check that $\mathbf{EA} = \mathbf{I}$ and also $\mathbf{Eh}$ gives the solution for x, y and z .

At present this simply appears as a viable but somewhat tedious method of finding an inverse – but now understood as a process it can be presented differently and ‘speeded-up’.

Our task is to find the inverse of $\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ 3 & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix}$. We did this by

using matrices $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \mathbf{E}_4$ and $\mathbf{E}_5$ which eventually changed the matrix to $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The effect of the $\mathbf{E}$ matrices is to perform row

reductions. We can compute the effect of the $\mathbf{E}$ matrices as we proceed so that at the end we shall not have to multiply them out to obtain the inverse. To be able to start the computation we use the matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$(1) \quad \begin{array}{ccc|ccc} & \mathbf{A} & & & \mathbf{I} & & \\ \begin{pmatrix} 1 & -1 & 2 \\ 3 & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix} & & & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array}.$$

$$\text{Now } \mathbf{E}_1 \mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ 3 & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 4 & -3 \end{pmatrix} \quad \text{and}$$

$$\mathbf{E}_1 \mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}. \text{ We may put together the results of using } \mathbf{E}_1 \text{ on } \mathbf{A}$$

and $\mathbf{I}$ as

$$(2) \quad \begin{array}{ccc|ccc} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 4 & -3 \end{pmatrix} & & \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \end{array}.$$

The effect of using $\mathbf{E}_1$ on $\mathbf{A}$ and $\mathbf{I}$ is to perform row reductions, and these reductions are performed just as we did originally on the equations. In (1) row 1 times -3 to add to row 2, row 1 multiplied by -1 to add to row 3 to give rows 2 and 3 in (2).

Proceeding to use $\mathbf{E}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix}$ with $\mathbf{E}_1\mathbf{A}$ and $\mathbf{E}_1\mathbf{I}$ we obtain

$$\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 17 \end{pmatrix} \text{ and } \mathbf{E}_2\mathbf{E}_1\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 11 & -4 & 1 \end{pmatrix}. \text{ Putting these}$$

results together as

$$(3) \left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -5 & -3 & 1 & 0 \\ 0 & 0 & 17 & 11 & -4 & 1 \end{array} \right)$$

we see row reduction has again been used. The effect of $\mathbf{E}_2$ is to take row 2 in (2), multiply it by -4 and add it to row 3 to give row 3 in (3). By doing this reduction to the whole row in (2) we are computing $\mathbf{E}_2\mathbf{E}_1\mathbf{I}$ on the right-hand side and hence the $\mathbf{E}$ matrices giving the inverse of $\mathbf{A}$ are being computed as we proceed.

$\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}$ and $\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{I}$ are computed by dividing row 3 in (3) throughout by 17, to give row 3 in (4).

$$(4) \left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -5 & -3 & 1 & 0 \\ 0 & 0 & 1 & \frac{11}{17} & -\frac{4}{17} & \frac{1}{17} \end{array} \right).$$

$\mathbf{E}_4$ is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}$, which has the effect of taking row 3 of (4) multi-

plying it by 5 and adding it to row 2 to give row 2 in (5). Thus for (5) we have $\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}$ and $\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{I}$ as

$$(5) \left(\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{4}{17} & -\frac{3}{17} & \frac{5}{17} \\ 0 & 0 & 1 & \frac{11}{17} & -\frac{4}{17} & \frac{1}{17} \end{array} \right).$$

And for $\mathbf{E}_5 = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ which adds row 1 to row 2 and multiplies

row 3 by -2 and adds it to row 1 to give rows 1 and 2 in (6), giving $\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{A}$ and $\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1\mathbf{I}$ computed as

$$(6) \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{17} & \frac{5}{17} & \frac{3}{17} \\ 0 & 1 & 0 & \frac{4}{17} & -\frac{3}{17} & \frac{5}{17} \\ 0 & 0 & 1 & \frac{11}{17} & -\frac{4}{17} & \frac{1}{17} \end{array} \right).$$

We have the inverse of $\mathbf{A}$ on the right-hand side. The clue to the process is the changing of $\mathbf{A}$ by row reduction into the identity matrix, performing the same reductions on the right-hand side. Obtaining the identity matrix on the left-hand side is a systematic process which is as follows:

If $\mathbf{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ then

- (i) obtain a 1 in place a – divide top row throughout by a
- (ii) obtain zeros on places d and g – multiply row 1 appropriately and add
- (iii) obtain a 1 in place e
- (iv) obtain zeros in places b and h
- (v) obtain 1 in place i
- (vi) obtain zeros in places c and f .

(This reduces some of the work we have done above – more could be done in (3) and (5) to obtain zeros.)

Example

To find the inverse of $\mathbf{A} = \begin{pmatrix} 2 & 1 & -3 \\ 4 & -1 & -2 \\ -1 & 1 & 2 \end{pmatrix}$

$\mathbf{A}$	$\mathbf{I}$	Instructions for working from one set to the next
$\left(\begin{array}{ccc ccc} 2 & 1 & -3 & 1 & 0 & 0 \\ 4 & -1 & -2 & 0 & 1 & 0 \\ -1 & 1 & 2 & 0 & 0 & 1 \end{array} \right)$		Divide row 1 by 2
$\left(\begin{array}{ccc ccc} 1 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 4 & -1 & -2 & 0 & 1 & 0 \\ -1 & 1 & 2 & 0 & 0 & 1 \end{array} \right)$		row 1 $\times -4$ add to row 2 row 1 $\times 1$, add to row 3
$\left(\begin{array}{ccc ccc} 1 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & -3 & 4 & -2 & 1 & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{array} \right)$		divide row 2 by -3
$\left(\begin{array}{ccc ccc} 1 & \frac{1}{2} & -\frac{3}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -\frac{4}{3} & -\frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{array} \right)$		row 2 $\times -\frac{1}{2}$ add to row 1 row 2 $\times -\frac{3}{2}$ add to row 3
$\left(\begin{array}{ccc ccc} 1 & 0 & -\frac{5}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 1 & -\frac{4}{3} & -\frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{5}{2} & -\frac{1}{2} & \frac{1}{2} & 1 \end{array} \right)$		divide row 3 by $\frac{5}{2}$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -\frac{5}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\ 0 & 1 & -\frac{4}{3} & \frac{2}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{1}{5} & \frac{1}{5} & \frac{2}{5} \end{array} \right) \quad \begin{array}{l} \text{row } 3 \times \frac{4}{3} \text{ add to row } 2 \\ \text{row } 3 \times \frac{5}{6} \text{ add to row } 1 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & \frac{2}{5} & -\frac{1}{15} & \frac{8}{15} \\ 0 & 0 & 1 & -\frac{1}{5} & \frac{1}{5} & \frac{2}{5} \end{array} \right)$$

This is still an arithmetical burden, but it presents a systematic process (which can be programmed for a computer) of finding the inverse of any (square) matrix. Another method involving determinants is available for 3×3 matrices, but it is perhaps best to keep to the given standard method for all cases.

Exercises

4.4 Find the inverse of the matrices

$$\begin{array}{lll} \text{(i)} \begin{pmatrix} 1 & 0 & 3 \\ 1 & 1 & 0 \\ 2 & 1 & 4 \end{pmatrix} & \text{(ii)} \begin{pmatrix} 1 & -3 & 2 \\ 1 & -2 & 1 \\ 2 & -1 & 3 \end{pmatrix} & \text{(iii)} \begin{pmatrix} 3 & 0 & 3 \\ 2 & 1 & -1 \\ -1 & 3 & 2 \end{pmatrix} \\ \text{(iv)} \begin{pmatrix} 2 & 0 & 1 & -2 \\ 1 & 1 & 1 & -1 \\ 0 & -3 & -1 & -2 \\ 2 & -1 & 0 & 1 \end{pmatrix} & & \end{array}$$

4.5 Choose a 2×2 matrix and find its inverse by choosing the suitable **E** matrices. Now look at the geometrical aspect. Use your matrix to transform the unit square. With each of the **E** matrices show diagrammatically how they systematically restore the transformed figure back to the unit square.

5

THE EIGENVECTOR

5.1 Invariants in a mapping

In this section we shall consider what properties of a figure remain unchanged when a transformation is made. The matrix to rotate the plane through an angle α about O is $\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ i.e. if $P^*(x^*, y^*)$ is the new point and $P(x, y)$ the original point then

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

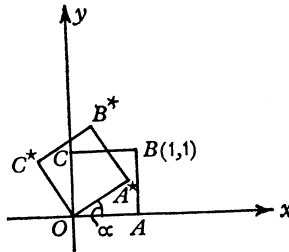


Fig. 5.1

Thus for the unit square in the diagram, Fig. 5.1

$$A(1, 0) \rightarrow A^*(\cos \alpha, \sin \alpha)$$

$$B(1, 1) \rightarrow B^*(\cos \alpha - \sin \alpha, \cos \alpha + \sin \alpha)$$

$$C(0, 1) \rightarrow C^*(-\sin \alpha, \cos \alpha).$$

Although the points A, B, C have moved in the transformation the figure formed by $OABC$ is still a square; the lengths of OA, BA etc. and the angles between the lines remain unaltered. So the expressions for these entities in terms of the new coordinates must be expressions which yield the same value as they did in the old coordinates. For example, $OA^* = \sqrt{(\cos^2 \alpha + \sin^2 \alpha)} = 1 = OA$. These are called *invariants* of the mapping.

Exercise

5.1 Find the length B^*C^* and the angle, $\angle OA^*B^*$ in the new co-ordinates and show that their values are the same as in the old co-ordinates.

On the other hand, the *direction* of the line OA is represented by the vector $(1, 0)$ in the old position and by the vector $(\cos \alpha, \sin \alpha)$ in the new position; so direction is not an invariant.

Does the matrix $\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ provide any other invariants in the geometry of its mapping? What of area? The area of the square remains the same, so here is another invariant.

If we consider the general matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, can we find the invariants of its mapping?

Again consider the unit square $OABC$.

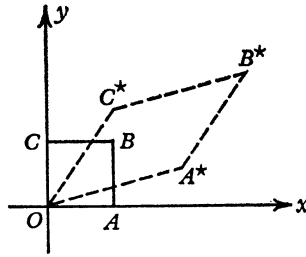


Fig. 5.2

Under this transformation (Fig. 5.2)

$$O \rightarrow O$$

$$A(1, 0) \rightarrow A^*(a, c)$$

$$B(1, 1) \rightarrow B^*(a+b, c+d)$$

$$C(0, 1) \rightarrow C^*(b, d).$$

Lengths and angles are now *not* invariant. O remains at O . What of area? – the area of $OA^*B^*C^*$ is not the same as that of $OABC$, so area is not an invariant.

Exercises

5.2 (a) Using the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ on the unit square $OABC$, show that the transformed figure $OA^*B^*C^*$ is a parallelogram.

(b) Show the area of the parallelogram $OA^*B^*C^*$ is $(ad - bc)$ (the determinant of the matrix). (Express the area as the sum or difference of triangles and trapezia as in Fig. 5.3.)

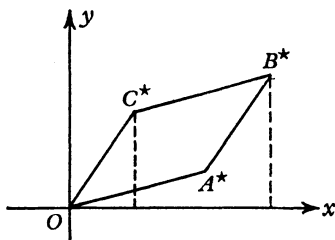


Fig. 5.3

(c) Deduce that the condition for the area to be an *invariant* of the mapping is $(ad - bc) = \pm 1$.

5.3 (a) For the following matrices, find the value of their determinants

(i) $\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$ (ii) $\begin{pmatrix} 3 & 2 \\ 5 & 4 \end{pmatrix}$ (iii) $\begin{pmatrix} 2 & 4 \\ 3 & 6 \end{pmatrix}$ (iv) $\begin{pmatrix} 1 & 3 \\ 3 & 6 \end{pmatrix}$

(v) $\begin{pmatrix} \cos 2\alpha & -\sin 2\alpha \\ \sin 2\alpha & \cos 2\alpha \end{pmatrix}$

(b) Which will have area as an invariant of the mapping?

(c) Show geometrically the effect on the unit square of the mapping represented by these matrices, explaining the relationship between the determinant of the matrix and the change in area produced by the mapping. Note that the determinant gives the *numerical* value of the area. A minus sign implies a reverse order of the vertices, i.e. a reflection in the mapping.

5.2 The eigenvector as an invariant of a matrix mapping

Besides lengths, angles and area, is there any other invariant which a matrix mapping may have? The following investigation leads to an important invariant possessed by all mappings produced by matrices.

Consider the effect of using the matrix $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix}$ to multiply the vector $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$: we have

$$\begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

so that the result is to multiply $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ by the scalar 2.

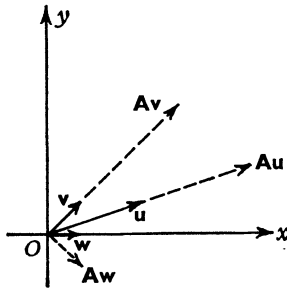


Fig. 5.4

Something similar happens if we use the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$;

$$\begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

This is an interesting and possibly significant situation, particularly when we consider the representation of the vectors as displacements, from the origin, in a plane, and the matrix as a transformation of the vectors. The situation is then that each of these vectors ($\mathbf{u}$ and $\mathbf{v}$ in Fig. 5.4) is transformed into a vector in the same direction as itself; $\mathbf{u} \rightarrow 2\mathbf{u}$ and $\mathbf{v} \rightarrow 4\mathbf{v}$.

The question now arises whether the matrix has the same effect on any other vectors, or whether $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ are special. Try $\mathbf{w} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, for example; this and its transform are shown in the diagram; they are not in the same direction.

Let us attempt systematically to find whether there are any other vectors $\begin{pmatrix} x \\ y \end{pmatrix}$ which behave with this matrix like $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. We must have

$$\begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{where } \lambda \text{ is a scalar}$$

$$\begin{cases} x + 3y = \lambda x \\ -x + 5y = \lambda y \end{cases}$$

$$\begin{cases} (1 - \lambda)x + 3y = 0 \\ -x + (5 - \lambda)y = 0 \end{cases} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (1).$$

From the consideration on page 37, if there are to be solutions to these equations other than $x = y = 0$ we must have

$$\det \begin{pmatrix} 1 - \lambda & 3 \\ -1 & 5 - \lambda \end{pmatrix} = 0.$$

Hence

$$\begin{aligned}(1-\lambda)(5-\lambda)+3 &= 0, \\ \lambda^2-6\lambda+8 &= 0, \\ (\lambda-2)(\lambda-4) &= 0 \\ \lambda &= 2 \text{ or } 4.\end{aligned}$$

(These are the two multiplying factors we found previously.)

To find whether any other vectors $\begin{pmatrix} x \\ y \end{pmatrix}$ exist we put each of these values of λ in turn, back into the equations (1). Do this, first with $\lambda=2$ in both equations. What do you notice about the two equations? And what values of $\begin{pmatrix} x \\ y \end{pmatrix}$ are solutions? Are there any others besides $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$? What about $\begin{pmatrix} 6 \\ 2 \end{pmatrix}$ or $\begin{pmatrix} 3k \\ k \end{pmatrix}$? Now put $\lambda=4$ into both equations and again find the solutions for $\begin{pmatrix} x \\ y \end{pmatrix}$. State your conclusions about the existence of further vectors which behave like $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

We now give these special vectors a name.

If $\mathbf{A}$ is any matrix, $\mathbf{x}$ a vector and λ a scalar, such that $\mathbf{Ax} = \lambda\mathbf{x}$, then $\mathbf{x}$ is called an *eigenvector* of $\mathbf{A}$, with *eigenvalue* λ .

Example

Find the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$.

Let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$ and let $\mathbf{x}$ be an eigenvector of $\mathbf{A}$ with eigenvalue λ , then $\mathbf{Ax} = \lambda\mathbf{x}$ and $\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$,

$$x + 2y = \lambda x$$

$$2x - 2y = \lambda y,$$

$$\text{and} \quad \left. \begin{aligned} (1-\lambda)x + 2y &= 0 \\ 2x + (-2-\lambda)y &= 0 \end{aligned} \right\} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (1).$$

For equations (1) to have solution other than $x=0=y$ then

$$\det \begin{pmatrix} 1-\lambda & 2 \\ 2 & -2-\lambda \end{pmatrix} = 0$$

$$(1-\lambda)(-2-\lambda) - 4 = 0,$$

$$\lambda^2 + \lambda - 6 = 0,$$

$$(\lambda-2)(\lambda+3) = 0$$

therefore $\lambda=2$ or -3 , the eigenvalues.

When $\lambda=2$, the equation (1) becomes $-x+2y=0$ and $2x-4y=0$, both equations being the same; $x-2y=0$. This is the eigenvector line corresponding to the eigenvalue of 2. Particular eigenvectors $\begin{pmatrix} x \\ y \end{pmatrix}$ are found by finding values of x and y which satisfy $x-2y=0$; e.g. $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$; $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$; etc. Note that for any such vector, say $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$, the matrix $\mathbf{A}$ has the effect of multiplying by

$$\lambda \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}; \quad \text{thus} \quad \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix} = 2 \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

When $\lambda = -3$, the equations (1) become $4x+2y=0$ and $2x+y=0$. Hence the eigenvector line is $2x+y=0$, with an eigenvector $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

5.3 Finding the eigenvectors of a matrix

We must summarise the above work by dealing with the general case.

The eigenvectors of the matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ are found by considering the mapping of $P(x, y)$ onto $P^*(\lambda x, \lambda y)$; that is $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$. The set of linear equations giving this mapping are

$$\begin{cases} (a-\lambda)x + by = 0 \\ cx + (d-\lambda)y = 0 \end{cases} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (1)$$

or $\mathbf{A}_1 \mathbf{x} = \mathbf{0}$, where $\mathbf{A}_1 = \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}$.

The condition that equations (1) have solutions other than $x=0=y$ is $\det \mathbf{A}_1 = 0$; that is

$$\det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = 0$$

or $(a-\lambda)(d-\lambda) - bc = 0 \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (2).$

Equation (2) is known as the *characteristic equation* of the matrix. The values of λ from (2) are substituted into (1) to give the eigenvector line – an equation of a line through the origin such that any point on the line is mapped (by the matrix) onto another point on the same line.

Since

$$\mathbf{A}_1 = \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{A} - \lambda \mathbf{I}$$

the characteristic equation can be written as $\det (\mathbf{A} - \lambda \mathbf{I}) = 0$.

Exercises

- 5.4** Find the eigenvalues, eigenvector lines and eigenvectors of the matrices

$$(i) \begin{pmatrix} 5 & 4 \\ 1 & 2 \end{pmatrix} \quad (ii) \begin{pmatrix} 6 & 2 \\ 2 & 9 \end{pmatrix}$$

- 5.5** Find the eigenvalues of $\begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 2 & 5 \\ -1 & -2 \end{pmatrix}$.

- 5.6** Write and simplify the characteristic equation of the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Deduce what you can about the sum and product of the eigenvalues.

- 5.7** Is it possible for a real matrix to have complex eigenvalues? Try to make one up.

- 5.8** Is it possible for a real matrix to have a single repeated eigenvalue? Try to make up some which have.

- 5.9** What can you say about the eigenvalues of a diagonal matrix? (One with zeros everywhere except on the principal diagonal; write down one and try it.)

- 5.10** What are the eigenvalues of a matrix of the form $k\mathbf{I}$?

- 5.11** What are the eigenvalues of the matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$?

- 5.12** Find and note the eigenvalues of the following matrices:

$$\begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 1 & 6 \end{pmatrix}; \quad \begin{pmatrix} 4 & 5 \\ 6 & -1 \end{pmatrix} \begin{pmatrix} 4 & 9 \\ 4 & -1 \end{pmatrix}.$$

- 5.13** Find the eigenvector lines of the matrices

$$\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}, \quad \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 5 \\ -1 & -2 \end{pmatrix}.$$

- 5.14** Investigate such further matrices as you need to, to draw conclusions about the eigenvectors of (a) matrices with real, distinct eigenvalues; (b) matrices with complex eigenvalues; (c) matrices of the form $k\mathbf{I}$; (d) matrices with a single repeated eigenvalue; (e) diagonal matrices.

5.4 The eigenvectors of a 3×3 matrix

Let $\mathbf{A} = \begin{pmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix}$, then the eigenvector of $\mathbf{A}$, $\mathbf{x}$ is given when

$\mathbf{Ax} = \lambda \mathbf{x}$, where λ is the corresponding eigenvalue. With $\mathbf{A}$ as a 2×2 matrix, the eigenvector is two-dimensional; now $\mathbf{A}$ is a 3×3 matrix and $\mathbf{x}$ is a three dimensional vector.

Hence

$$\begin{pmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

$$\begin{aligned} 2x - 2y + 3z &= \lambda x \\ x + y + z &= \lambda y \\ x + 3y - z &= \lambda z, \end{aligned}$$

$$\left. \begin{aligned} (2 - \lambda)x - 2y + 3z &= 0 \\ x + (1 - \lambda)y + z &= 0 \\ x + 3y + (-1 - \lambda)z &= 0 \end{aligned} \right\} \quad . \quad . \quad . \quad (1).$$

As before, equations (1) have solutions other than $x = 0 = y = z$ if

$$\det \begin{pmatrix} 2 - \lambda & -2 & 3 \\ 1 & 1 - \lambda & 1 \\ 1 & 3 & -1 - \lambda \end{pmatrix} = 0.$$

The cubic equation for λ , the characteristic equation for $\mathbf{A}$, may be obtained by expanding the third order determinant, or we may obtain it by eliminating x , from equations (1) and then using a second order determinant. This latter process is as follows

$$(2 - \lambda)x - 2y + 3z = 0 \quad . \quad . \quad . \quad (a)$$

$$x + (1 - \lambda)y + z = 0 \quad . \quad . \quad . \quad (b)$$

$$x + 3y + (-1 - \lambda)z = 0 \quad . \quad . \quad . \quad (c)$$

$(2 - \lambda)(b) - (a)$ gives

$$(1 - \lambda)(2 - \lambda)y + 2y + (2 - \lambda)z - 3z = 0$$

that is

$$(\lambda^2 - 3\lambda + 4)y + (-1 - \lambda)z = 0 \quad . \quad . \quad . \quad (d)$$

$(b) - (c)$ gives

$$(-2 - \lambda)y + (2 + \lambda)z = 0 \quad . \quad . \quad . \quad (e)$$

From equations (d) and (e) we have solutions other than $y=0=z$ when $\det \begin{pmatrix} \lambda^2 - 3\lambda + 4 & -1 - \lambda \\ -2 - \lambda & 2 + \lambda \end{pmatrix} = 0$

$$\begin{aligned} (\lambda^2 - 3\lambda + 4)(2 + \lambda) - (-2 - \lambda)(-1 - \lambda) &= 0, \\ \lambda^3 - 2\lambda^2 - 5\lambda + 6 &= 0, \\ (\lambda - 1)(\lambda + 2)(\lambda - 3) &= 0. \end{aligned}$$

The eigenvalues of $\mathbf{A}$ are 1, -2, and 3.

When $\lambda = 1$, in equations (1)

$$\begin{aligned} x - 2y + 3z &= 0 & . & . & . & . & \text{(i),} \\ x &+ z &= 0 & . & . & . & \text{(ii),} \\ x + 3y - 2z &= 0 & . & . & . & . & \text{(iii).} \end{aligned}$$

From (ii) $x = -z$, so (i) becomes $-2y + 2z = 0$ and (iii) becomes $3y - 3z = 0$.

Both of these equations give $y = z$.

The eigenvector line is given by $x/-1 = y/1 = z/1$, which is a line through the origin. A particular eigenvector is $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$. Note that

$$\begin{pmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = 1 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

When $\lambda = -2$, the eigenvector line is $x/11 = y/1 = z/-14$, an eigenvector is $\begin{pmatrix} 11 \\ 1 \\ -14 \end{pmatrix}$; and when $\lambda = 3$ the eigenvector line is $x/1 = y/1 = z/1$.

Exercise

5.15 Find the eigenvalues and eigenvectors of the matrices

$$\text{(i)} \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} \quad \text{(ii)} \begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

5.5 Transpose of a matrix

The *transpose* $\mathbf{A}'$ of a matrix $\mathbf{A}$ is the matrix formed from $\mathbf{A}$ by writing the rows of $\mathbf{A}$ as columns; the n th row of $\mathbf{A}$ becomes the n th column of $\mathbf{A}'$.

Thus the transpose ($\mathbf{A}'$) of the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ is

$$\mathbf{A}' = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

If a matrix is square its transpose is a reflection in the leading diagonal;

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \mathbf{A}' = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}.$$

5.6 Transpose of a product

Exercise

5.16 $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} -2 & 5 \\ 3 & 6 \end{pmatrix}$ find $\mathbf{AB}$, $\mathbf{A'B'}$, $\mathbf{B'A'}$, $(\mathbf{AB})'$.

Is there any relationship between $\mathbf{A'}$, $\mathbf{B'}$ and $(\mathbf{AB})'$?

Using 2×2 general matrices, prove the relationship between $\mathbf{A'}$, $\mathbf{B'}$ and $(\mathbf{AB})'$ is $(\mathbf{AB})' = \mathbf{B'A'}$.

5.7 Orthogonal vectors

If $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ is a column vector, then $\mathbf{u}' = (u_1, u_2)$ is a row vector. When $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ the product $\mathbf{u'v}$ is $(u_1, u_2) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = (u_1v_1 + u_2v_2)$ – a scalar; in fact the ‘inner product’ $\mathbf{u} \cdot \mathbf{v}$ of the vectors $\mathbf{u}$, $\mathbf{v}$. If $\mathbf{u} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 6 \\ -2 \end{pmatrix}$ then $\mathbf{u' \cdot v} = 0$, and as the gradient of $\mathbf{u}$ is 3 and that of $\mathbf{v}$ is $-\frac{1}{3}$, $\mathbf{u}$ and $\mathbf{v}$ are perpendicular. If $\mathbf{u' \cdot v} = 0$ the vectors $\mathbf{u}$ and $\mathbf{v}$ will be perpendicular or *orthogonal*.

5.8 Symmetric matrices

If a matrix $\mathbf{A}$ is such that $\mathbf{A} = \mathbf{A'}$, then $\mathbf{A}$ is said to be *symmetric*. A symmetric matrix must be square.

Exercise

5.17 Find the eigenvalues and show that the eigenvector lines of the symmetric matrix $\begin{pmatrix} 1 & 4 \\ 4 & -5 \end{pmatrix}$ are orthogonal. Show that the general symmetric matrix $\mathbf{M} = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ always has real and distinct eigenvalues and orthogonal eigenvectors.

5.9 Inverse of a matrix – another method

The characteristic equation of the matrix $\mathbf{A} = \begin{pmatrix} 3 & 4 \\ 2 & 5 \end{pmatrix}$ can be verified as $\lambda^2 - 8\lambda + 7 = 0$. Replace λ by $\mathbf{A}$, then

$$\begin{aligned} \mathbf{A}^2 - 8\mathbf{A} + 7\mathbf{I} &= \begin{pmatrix} 3 & 4 \\ 2 & 5 \end{pmatrix}^2 - 8\begin{pmatrix} 3 & 4 \\ 2 & 5 \end{pmatrix} + 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 17 & 32 \\ 16 & 33 \end{pmatrix} - \begin{pmatrix} 24 & 32 \\ 16 & 40 \end{pmatrix} + \begin{pmatrix} 7 & 0 \\ 0 & 7 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \end{aligned}$$

$\mathbf{A}$ satisfies its own characteristic equation.

The inverse $\mathbf{A}^{-1}$ of $\mathbf{A}$ can now be found as follows: $\mathbf{A}^2 - 8\mathbf{A} = -7\mathbf{I}$, $\mathbf{A}(\mathbf{A} - 8\mathbf{I}) = -7\mathbf{I}$ (Distributive Law).

Premultiply each side by $\mathbf{A}^{-1}$, then $\mathbf{A}^{-1}\mathbf{A}(\mathbf{A} - 8\mathbf{I}) = \mathbf{A}^{-1} \cdot -7\mathbf{I}$, that is

$$\mathbf{A} - 8\mathbf{I} = -7\mathbf{A}^{-1}$$

hence

$$\begin{aligned} \mathbf{A}^{-1} &= -\frac{1}{7}(\mathbf{A} - 8\mathbf{I}) = -\frac{1}{7}\left[\begin{pmatrix} 3 & 4 \\ 2 & 5 \end{pmatrix} - \begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix}\right] \\ &= -\frac{1}{7}\begin{pmatrix} -5 & 4 \\ 2 & -3 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 5 & -4 \\ -2 & 3 \end{pmatrix}. \end{aligned}$$

The crux of this method is the fact that the matrix $\mathbf{A}$ satisfies its own characteristic equation. Will this always be true? Write down other 2×2 matrices (or the general 2×2 matrix) and see if they satisfy their characteristic equations; then find their inverses by this method. It is possible by this process to find the inverse of 3×3 matrices. (The general theorem that every square matrix satisfies its characteristic equation is proved in Birkhoff and MacLane, *A Survey of Modern Algebra*, 1953 edition, p. 320.)

Exercise

5.18 Show that the characteristic equation of the matrix

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 3 \\ 1 & 1 & 0 \\ 2 & 1 & 4 \end{pmatrix}$$

is $\lambda^3 - 6\lambda^2 + 3\lambda - 1 = 0$. Replace λ by $\mathbf{M}$ and establish that

$$\mathbf{M}^3 - 6\mathbf{M}^2 + 3\mathbf{M} - \mathbf{I} = 0.$$

Multiply through by $\mathbf{M}^{-1}$, showing $\mathbf{M}^{-1} = \mathbf{M}^2 - 6\mathbf{M} + 3\mathbf{I}$ and hence find $\mathbf{M}^{-1}$.

6

APPLICATIONS OF EIGENVECTORS TO LINEAR MAPPINGS

6.1 Introduction

This section consists of an extended investigation in which eigenvectors play the crucial role. It is not, however, essential for understanding the later parts of the book.

The investigation consists of trying to discover what transformation of the plane a given matrix will perform. It may be an enlargement, reflection, rotation, shear or projection or combination of these.

The general basis of the method will be to rotate the axes of co-ordinates so that they lie along the axes of the transformation (for example in the case of a reflection, along and perpendicular to the mirror), thus reducing the matrix of the transformation to an easily recognisable form. In seeking the axes of the transformation, the eigenvector of the matrix will have an important role.

6.2 Transformation matrices and their eigenvectors

The reflection matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, which reflects points in the x axis, leaves points on the x axis where they are and reflects points of the y axis onto other points of the y axis. So the axes are eigenvectors of this reflection matrix; and since $(x, 0) \rightarrow (x, 0)$, while $(0, y) \rightarrow (0, -y)$, we see that the eigenvalues are 1 and -1 respectively for the two eigenvectors.

The relation between the different kinds of mapping and their eigenvectors is shown in Table 6.1 (p. 56).

6.3 To find the mapping of a given matrix

We shall now consider how far this set is sufficient for expressing any given mapping. We may note straight away that we do not need all three of $\mathbf{D}$, $\mathbf{E}$ and $\mathbf{L}$, since by taking $k=b$ and $l=a/b$, we have $\mathbf{D}=\mathbf{EL}=\mathbf{LE}$.

As a first example of what we propose to do, consider the matrix $\mathbf{M} = \begin{pmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{pmatrix}$. Its eigenvalues are readily shown to be $+1$,

Matrix	description	deter- minant	eigen- values	eigenvectors
(a) $\mathbf{D} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$	double stretch	ab	a, b	x axis, y axis
(b) $\mathbf{F} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	reflection in x axis	-1	$1, -1$	x axis, y axis
(c) $\mathbf{E} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$	enlargement	k^2	k	any vector
(d) $\mathbf{L} = \begin{pmatrix} l & 0 \\ 0 & 1 \end{pmatrix}$	linear stretch in x axis direction	l	$l, 1$	x axis, y axis
(e) $\mathbf{S} = \begin{pmatrix} 1 & S \\ 0 & 1 \end{pmatrix}$	shear in x axis direction	1	1	x axis
(f) $\mathbf{P} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	projection parallel y axis into x axis	0	$1, 0$	x axis, y axis
(g) $\mathbf{R} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$	rotation of points through α	1	$\cos \alpha$ $\pm i \sin \alpha$	not real

(a, b, k, l are >0)

Table 6.1

-1 and its eigenvectors $\begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix}$ and $\begin{pmatrix} -\sin \beta \\ \cos \beta \end{pmatrix}$. (Note that they are orthogonal and at an angle β to the original axes.) This suggests that $\mathbf{M}$ represents a reflection in the line making an angle β with the x axis, which can be verified directly but which will also be proved by what follows here. Suppose we now consider a *rotation of axes* through an angle β ; the reflection is now in the new x axis, so the matrix $\mathbf{M}$, when transformed in the appropriate way, must become $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, which is one of the basic set we have chosen. It remains to decide what is the appropriate way to transform the matrix of a mapping when the axes of coordinates are rotated. Figure 6.1 shows original axes (x, y) , new axes (x_1, y_1) rotated through an angle β from the old, and a typical point P which maps into P^* . The coordinates of P and P^* may be expressed with respect to either the old or the new axes. If the matrix of the mapping is $\mathbf{M}$ with respect to the old axes, we have writing $\mathbf{x}$ for $\begin{pmatrix} x \\ y \end{pmatrix}$ and $\mathbf{x}_1$ for $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$

$$\mathbf{x}^* = \mathbf{M}\mathbf{x} \quad . \quad . \quad . \quad . \quad (1).$$

The relation between the old and new coordinates of P and P^* is given by

$$\mathbf{x}_1 = \mathbf{R}\mathbf{x} \quad . \quad . \quad . \quad . \quad (2)$$

and

$$\mathbf{x}_1^* = \mathbf{R}\mathbf{x}^* \quad . \quad . \quad . \quad . \quad (3)$$

where $\mathbf{R} = \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix}$ – the matrix for expressing the change in the coordinates of points when the axes are rotated through $+\beta$, see p. 15.

Thus we have

$$\begin{aligned} \mathbf{x}_1^* &= \mathbf{R}\mathbf{x}^* && \text{from equation (3)} \\ &= \mathbf{R}\mathbf{M}\mathbf{x} && \text{from equation (1)} \\ &= \mathbf{R}\mathbf{M}(\mathbf{R}^{-1}\mathbf{x}_1) && \text{from equation (2)} \\ &= (\mathbf{R}\mathbf{M}\mathbf{R}^{-1})\mathbf{x}_1 \end{aligned}$$

which shows that the matrix giving the same mapping as $\mathbf{M}$ but with respect to the new axes is $\mathbf{R}\mathbf{M}\mathbf{R}^{-1}$.

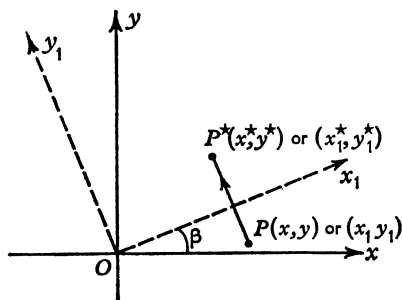


Fig. 6.1

Since this process is very important in the remainder of this section, and is not easy to understand, we give a second explanation of it which may help those who find the above difficult.

We are still referring to Fig. 6.1.

We need to rotate the *axes* through β into the x_1, y_1 , axes position, keeping the points fixed but referring them to axes rotated through β . If we consider a general point P in this mapping, then under a matrix $\mathbf{M}$, $P \rightarrow P^*$ where $\mathbf{x}^* = \mathbf{M}\mathbf{x}$. The old points in the plane, when the axes are rotated through an angle β , will become a set of points with respect to the new axes x_1, y_1 given by $\mathbf{x}_1 = \mathbf{R}\mathbf{x}$, where $\mathbf{R}$ is the *rotation of axes* matrix. Similarly the transformed points under the matrix $\mathbf{M}$, will have coordinates with respect to the x_1, y_1 axes given by $\mathbf{x}_1^* = \mathbf{R}\mathbf{x}^*$.

The transformed points taken *with respect to the rotated axes* are given by

$$\begin{aligned} \mathbf{x}_1^* &= \mathbf{R}\mathbf{x}^* \\ &= \mathbf{R}(\mathbf{M}\mathbf{x}) \quad (\text{from } \mathbf{x}^* = \mathbf{M}\mathbf{x}) \\ &= \mathbf{R}\mathbf{M}(\mathbf{R}^{-1}\mathbf{x}_1) \quad (\text{as } \mathbf{x}_1 = \mathbf{R}\mathbf{x} \text{ so that } \mathbf{x} = \mathbf{R}^{-1}\mathbf{x}_1) \\ &= \mathbf{R}\mathbf{M}\mathbf{R}^{-1}\mathbf{x}_1. \end{aligned}$$

So the effect of rotating the axes is that the matrix of the mapping changes from $\mathbf{M}$ to $\mathbf{R}\mathbf{M}\mathbf{R}^{-1}$ with respect to the new axes.

We now apply this result $\mathbf{M} \rightarrow \mathbf{RMR}^{-1}$ to the matrix we are considering,

$$\begin{pmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{pmatrix};$$

we have

$$\begin{aligned} \mathbf{RMR}^{-1} &= \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} \cos 2\beta & \sin 2\beta \\ \sin 2\beta & -\cos 2\beta \end{pmatrix} \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} \\ &= \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} \\ &\quad \times \begin{pmatrix} \cos 2\beta \cos \beta + \sin 2\beta \sin \beta & -\cos 2\beta \sin \beta + \sin 2\beta \cos \beta \\ \sin 2\beta \cos \beta - \cos 2\beta \sin \beta & -\sin 2\beta \sin \beta - \cos 2\beta \cos \beta \end{pmatrix} \\ &= \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} \cos \beta & \sin \beta \\ \sin \beta & -\cos \beta \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \mathbf{F} \text{ as named in the list on p. 56.} \end{aligned}$$

Thus we have proved that when the axes are rotated through an angle β , the mapping produced by the matrix $\mathbf{M}$ becomes a reflection in the x_1 axis; so with respect to the old axes it must have been a reflection in the line making an angle β with the x axis.

If we look at this last relation from another point of view, and transform as follows:

$$\begin{aligned} \mathbf{RMR}^{-1} &= \mathbf{F} \\ \mathbf{MR}^{-1} &= \mathbf{R}^{-1}\mathbf{F}, \\ \mathbf{MR}^{-1}\mathbf{R} &= \mathbf{R}^{-1}\mathbf{FR}, \\ \mathbf{M} &= \mathbf{R}^{-1}\mathbf{FR}. \end{aligned}$$

We have achieved our aim of expressing $\mathbf{M}$ as a combination of matrices chosen from the basic set.

6.4 Another mapping

As a second example, consider the mapping produced by the matrix $\mathbf{K} = \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix}$. To transform this into a combination of matrices of the basic set, we have to find its eigenvectors and then rotate the axes of coordinates so that the new axes are in the directions of the eigenvectors. The eigenvalues turn out to be 4 and 2, with eigenvectors $\mathbf{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ respectively. Figure 6.2 shows the unit square and

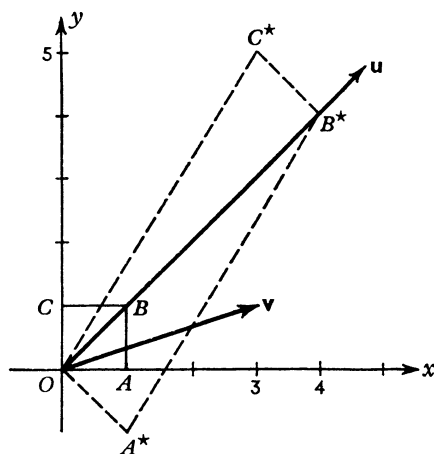


Fig. 6.2

the parallelogram into which it maps, and the two eigenvectors of the mapping.

We now choose one of the eigenvectors – it does not matter which – to become the new x axis (x_1). We choose $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. This is at 45° with the old x axis so we choose axes x_1, y_1 at 45° to the original x, y axes and use the matrix for rotating axes through 45° .

The matrix $\mathbf{R}$ required will be $\begin{pmatrix} \cos 45^\circ & \sin 45^\circ \\ -\sin 45^\circ & \cos 45^\circ \end{pmatrix}$, that is $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, and $\mathbf{R}^{-1}$ is $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.

The matrix representing the same mapping with respect to the new axes is given by

$$\begin{aligned}
 \mathbf{RMR}^{-1} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & 5 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\
 &= \frac{1}{2} \begin{pmatrix} 0 & 8 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\
 &= \frac{1}{2} \begin{pmatrix} 8 & 8 \\ 0 & 4 \end{pmatrix} \\
 &= \begin{pmatrix} 4 & 4 \\ 0 & 2 \end{pmatrix} \\
 &= \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}
 \end{aligned}$$

which is a shear of factor 1 followed by a double stretch of factors 4 and 2; these mappings being taken with respect to the x_1, y_1 axes. Thus this analysis of the matrix has enabled us to determine the nature of the mapping.

If we had taken the eigenvector $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ as an axis a similar result ensues.

6.5 Further examples of mappings

There now follow three worked examples on finding the nature of the mapping produced by a matrix. In these, various aspects of using the eigenvectors as axes have been developed. The chapter ends with a set of miscellaneous examples for the reader.

$$(i) \mathbf{M} = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}.$$

The characteristic equation of $\mathbf{M}$ is $(1-\lambda)(-1-\lambda)=0$, from which $\lambda = \pm 1$. The equations to give the eigenvectors are

$$(1-\lambda)x - 2y = 0 \quad . \quad . \quad . \quad (1)$$

and

$$0 \cdot x + (-1-\lambda)y = 0 \quad . \quad . \quad . \quad (2).$$

When $\lambda = 1$, from (1) $0 \cdot x - 2y = 0$, i.e. x can have any value and $y = 0$; which gives $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$; i.e. the x axis as an eigenvector. When $\lambda = -1$, from

$$(1) \quad 2x - 2y = 0, \text{ giving an eigenvector } \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Considering the eigenvector along the x axis, this implies $\mathbf{M}$ is already in basic form – since the basic matrices have the axes as their eigenvectors. So we test whether $\mathbf{M}$ is a direct combination of basic matrices.

By inspection $\mathbf{M} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ a reflection in the x axis followed by a shear of factor 2 in the x axis direction (Fig. 6.3). It is worthwhile to see how the eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ deals with the situation. Consider a new set of axes x_1, y_1 with the eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ as the new x_1 axis. The x_1 axis is thus at 45° to the x axis. The rotation of axes matrix

$$\mathbf{R} = \begin{pmatrix} \cos 45^\circ & \sin 45^\circ \\ -\sin 45^\circ & \cos 45^\circ \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

is needed to transform the coordinates of the points.

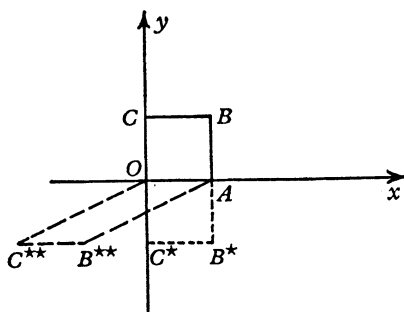


Fig. 6.3

The mapping of the matrix $\mathbf{M}$ when referred to the x_1y_1 axes is given by $\mathbf{RMR}^{-1}$ (refer to page 57)

$$\begin{aligned} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & -3 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} -2 & -4 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} -1 & -2 \\ 0 & 1 \end{pmatrix} = \mathbf{M}_1. \end{aligned}$$

$\mathbf{M}_1$ is the matrix now referred to the x_1, y_1 axes. We express this in terms of the basic matrices.

$\mathbf{M}_1 = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ – a reflection in the y_1 axis followed by a shear of factor -2 in the x_1 axis direction. The reader is invited to find the coordinates of the old unit square with respect to the x_1, y_1 axes and follow through a reflection in the y_1 axis and the shear $\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$ verifying that the same transformed figure is obtained.

$$(ii) \mathbf{M} = \begin{pmatrix} 2 & 1 \\ 5 & -2 \end{pmatrix}, \quad \det \mathbf{M} = -9.$$

The characteristic equation is $(2 - \lambda)(-2 - \lambda) - 5 = 0$ from which
 $\lambda^2 = 9$ and $\lambda = \pm 3$.

The equations to give the eigenvectors are

$$(2 - \lambda)x + y = 0 \quad . \quad . \quad . \quad (1)$$

$$\text{and} \quad 5x + (-2 - \lambda)y = 0 \quad . \quad . \quad . \quad (2).$$

When $\lambda = 3$, from (1) $-x + y = 0$, giving an eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and, when $\lambda = -3$, from (1) $5x + y = 0$, giving an eigenvector $\begin{pmatrix} 1 \\ -5 \end{pmatrix}$. Considering the eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ the rotation of axes matrix for referring the points of the mapping with respect to axes x_1, y_1 inclined at 45° to the x, y axes is $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.

Thus $\mathbf{M}$ referred to the new axes is given by

$$\begin{aligned} \mathbf{RMR}^{-1} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 5 & -2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -4 \\ 0 & -3 \end{pmatrix} = \mathbf{M}_1. \end{aligned}$$

By inspection

$$\begin{aligned} \mathbf{M}_1 &= \begin{pmatrix} 3 & 4 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= 3 \begin{pmatrix} 1 & \frac{4}{3} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

– a reflection in the x_1 axis, followed by a shear of factor $\frac{4}{3}$ in the x_1 axis direction and with an enlargement factor of 3.

If the eigenvector $\begin{pmatrix} 1 \\ -5 \end{pmatrix}$ is considered the rotation of axes matrix is $\frac{1}{\sqrt{26}} \begin{pmatrix} 1 & -5 \\ 5 & 1 \end{pmatrix}$ since $\tan \alpha = -5$. The reader is invited to check that the same transformation is produced (but with slightly different basic matrices) with this eigenvector as the new x_1 axis direction as was found with the eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ as the x_1 axis direction.

$$(iii) \mathbf{M} = \frac{1}{5} \begin{pmatrix} 11 & -2 \\ -2 & 14 \end{pmatrix}.$$

We can find the eigenvalues by considering $(11 - \lambda)(14 - \lambda) - 4 = 0$, from which $\lambda^2 - 25\lambda + 150 = 0$ and $\lambda = 10$ or 15 .

The equations to give the eigenvectors are

$$(11 - \lambda)x - 2y = 0 \quad . \quad . \quad . \quad (1)$$

$$\text{and} \quad -2x + (14 - \lambda)y = 0 \quad . \quad . \quad . \quad (2).$$

When $\lambda = 10$, from (1), $x - 2y = 0$, giving an eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

When $\lambda = 15$, from (1), $-4x - 2y = 0$, giving an eigenvector $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

Considering the eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, the rotation of axes matrix is $\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$ where $\tan \alpha = \frac{1}{2}$; i.e. $\mathbf{R} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$.

The transformed matrix of the mapping, when referred to the x_1y_1 axes when the direction of x_1 with respect to the x axis is $\tan^{-1}(\frac{1}{2})$ is given by $\mathbf{RMR}^{-1}$

$$\begin{aligned} &= \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 11 & -2 \\ -2 & 14 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \\ &= \frac{1}{25} \begin{pmatrix} 20 & 10 \\ -15 & 30 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \\ &= \frac{1}{25} \begin{pmatrix} 50 & 0 \\ 0 & 75 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \end{aligned}$$

– a double stretch of factors 2 and 3 with respect to the x_1 and y_1 axes respectively.

6.6 Classification of a mapping

We now begin to classify the mapping of a matrix by the consideration of its eigenvalues and eigenvectors, using exercises for the reader.

1. Show that the matrix $\frac{1}{5} \begin{pmatrix} 3 & 4 \\ -4 & -3 \end{pmatrix}$ gives a mapping in which points are reflected in the line $2y = x$. (Find the eigenvectors of the matrix and show by a suitable rotation of axes that the matrix mapping is a reflection in one of the new axes.)

Verify the determinant of the matrix is -1 .

Verify also that the following matrices have a determinant of -1 and find what mappings they produce.

$$(i) \frac{1}{5} \begin{pmatrix} 3 & -4 \\ -4 & -3 \end{pmatrix} \quad (ii) \frac{1}{25} \begin{pmatrix} -24 & -7 \\ -7 & 24 \end{pmatrix}.$$

2. Show that the matrix $\mathbf{M} = \begin{pmatrix} 4 & 5 \\ 5 & 4 \end{pmatrix}$ has determinant -9 (i.e. of the form $-k^2$). Verify it has eigenvalues of 9 and -1 and eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ respectively. By a rotation of axes through 45° corresponding to the eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ show the matrix transforms into the matrix $\mathbf{M}_1 = \begin{pmatrix} 9 & 0 \\ 0 & -1 \end{pmatrix}$ when referred to the new axes. Verify this matrix $\mathbf{M}_1$ is $\begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ – a reflection in the new x axis followed by a linear stretch of factor 9 in the new x axis direction.

A determinant of the form $-k^2$ indicates the presence of a reflection component in the mapping which becomes clear when the matrix has been referred to suitable axes.

Investigate the following matrices which have determinants of the form $-k^2$.

$$(i) \begin{pmatrix} 3 & 2 \\ 2 & 0 \end{pmatrix} \quad (ii) \begin{pmatrix} 1 & 4 \\ 2 & -1 \end{pmatrix}.$$

If we now stop for a moment and look at the results we have obtained, we may note that, after suitable rotation of the axes, all the mappings considered reduce to combinations of some or all of shear, double stretch and reflection; also that the reduced matrices all have a zero in the lower left-hand corner. This latter point is worth investigating to see what conditions are necessary for it. It happens when we make the x axis an eigenvector of the matrix. Considering the general matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, if the x axis, represented by the vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is an eigenvector we must have $(a - \lambda)1 + b \cdot 0 = 0$ and $c \cdot 1 + (d - \lambda) \cdot 0 = 0$. The second of these gives $c = 0$.

Conversely, if $c = 0$ in any matrix, these equations show that $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ will be an eigenvector with eigenvalue $\lambda = a$. So we have proved that

$$c = 0 \Leftrightarrow (1, 0) \text{ is an eigenvector.}$$

The matrices so far considered have all had real non-zero eigenvalues. We consider next cases in which the eigenvalues are complex, and therefore there are no real eigenvectors. In these cases the mapping contains a rotation component which must be separated out before the rest can be analysed.

6.7 Example with complex eigenvalues

Consider the matrix $\mathbf{M} = \begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}$. This has eigenvalues $1 \pm 2i$.

We therefore try to express $\mathbf{M}$ as $\mathbf{M}_1\mathbf{R}$ where $\mathbf{R}$ is a rotation matrix and $\mathbf{M}_1$ has real eigenvalues.

Taking $\mathbf{R} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$, since $\mathbf{M} = \mathbf{M}_1\mathbf{R} \Rightarrow \mathbf{M}_1 = \mathbf{M}\mathbf{R}^{-1}$ we have

$$\begin{aligned} \mathbf{M}_1 &= \begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \\ &= \begin{pmatrix} 3 \cos \alpha - 2 \sin \alpha & -3 \sin \alpha - 2 \cos \alpha \\ 4 \cos \alpha - \sin \alpha & -4 \sin \alpha - \cos \alpha \end{pmatrix} \end{aligned}$$

and α has now to be chosen to make the eigenvalues of $\mathbf{M}_1$ real.

The condition for this in the general case is that

$$\det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = 0$$

shall have real roots; that is $(a - \lambda)(d - \lambda) - bc = 0$ has real roots; that is $\lambda^2 - (a + d)\lambda + (ad - bc) = 0$ has real roots, for which the condition is

$$(a + d)^2 - 4(ad - bc) \geq 0.$$

This gives $a^2 + 2ad + d^2 - 4ad + 4bc \geq 0$, or

$$(a - d)^2 + 4bc \geq 0. \quad . \quad . \quad . \quad (4).$$

This condition is not difficult to satisfy – in fact, it being an inequality, when applied to $\mathbf{M}_1$ it will give a whole range of values of α . We may therefore consider whether a particular choice of α within the possible range may be advantageous. For example, if in (4) we have $c = 0$, the inequality is certainly satisfied, and we then have a matrix $\mathbf{M}_1$ with zero in the lower left hand corner which, as we have shown above, (p. 64) already has the x axis as an eigenvector. We do not need therefore, to perform any change of axes; $\mathbf{M}_1$ is already in a form in which it can be reduced to a combination of shear and linear stretch. The working is as follows:

Choose α so that $4 \cos \alpha - \sin \alpha = 0$; i.e.

$$\tan \alpha = 4, \quad \sin \alpha = \frac{4}{\sqrt{17}}, \quad \cos \alpha = \frac{1}{\sqrt{17}}.$$

$$\begin{aligned}
 \text{Then } \mathbf{M}_1 &= \frac{1}{\sqrt{17}} \begin{pmatrix} 3 & -8 & -12 & -2 \\ 0 & -16 & -1 & 0 \end{pmatrix} \\
 &= \frac{1}{\sqrt{17}} \begin{pmatrix} -5 & -14 \\ 0 & -17 \end{pmatrix} \\
 &= \frac{-1}{\sqrt{17}} \begin{pmatrix} 5 & 0 \\ 0 & 17 \end{pmatrix} \begin{pmatrix} 1 & 2.8 \\ 0 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ \sqrt{17} & \sqrt{17} \end{pmatrix} \begin{pmatrix} 1 & 2.8 \\ 0 & 1 \end{pmatrix}
 \end{aligned}$$

which is a shear of factor 2.8, a double stretch with factors $5/\sqrt{17}$ and $\sqrt{17}$, and a 180° rotation, in succession.

6.8 Matrices with zero eigenvalues

It remains to consider matrices which have one or more zero eigenvalues. Take, for example, the matrix $\begin{pmatrix} 3 & -1 \\ -6 & 2 \end{pmatrix}$, which has eigenvalues 0, 5, or $\begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix}$, which has the eigenvalue 0 only. It is left as an investigation for the reader to show that in these cases it is still possible to perform a rotation of axes to make the x axis an eigenvector, and to remove stretch components as in the examples above. The first type then reduces to one of the forms $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, according to whether the new x axis, belongs to the zero or the non-zero eigenvalue; these give the first two projections shown in Fig. 6.4.

The matrix with eigenvalue of zero only reduces to $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ which gives the two-stage projection shown in the third diagram of Fig. 6.4. The details are left for the reader to investigate.

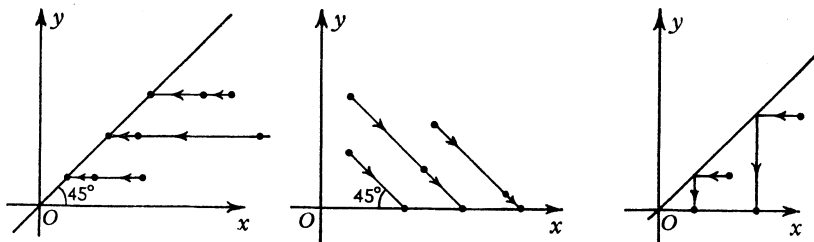


Fig. 6.4

6.9 Further investigations

1. For the reader who is familiar with the ideas of null-space and range of a linear mapping: relate these to the eigenvalues and eigenvectors of a singular matrix.

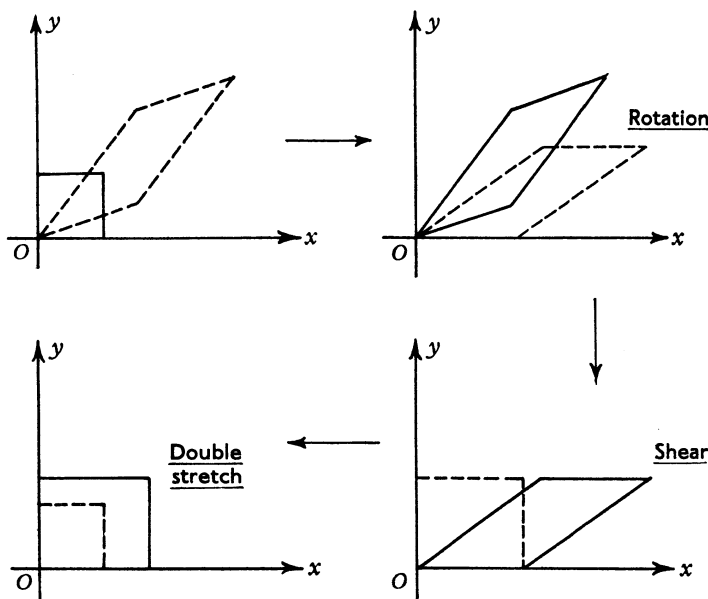


Fig. 6.5

2. As an alternative approach to the analysis of mappings produced by 2×2 matrices with non-zero determinant consider the unit square being mapped into a parallelogram, which can then be reduced to the square again by the sequence of operations shown in Fig. 6.5. Investigate whether, given the original matrix, the necessary matrices to perform this reduction can be found. As an extension, is this the only or the best possible order for the operations?

3. In the analysis in this section we have required the shear component to be in the direction of the x axis. If we allow it to be in any direction, can we eliminate the need for a *double stretch*, making do with a homogeneous enlargement $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ instead? If so, we can state that every non-singular linear mapping of the plane is equivalent, at most, to a shear and an enlargement, together with a rotation or a reflection. Investigate this.

NETWORKS AND COMMUNICATIONS

7.1 Introduction

In this chapter we consider the use of matrix methods in networks and communications. The application of matrices in this field is becoming extensive. The range of mathematical technique and difficulty is wide, but begins at such a level as to admit of matrix methods being taught through this topic in the classroom at the early secondary stage. Networks have in fact much to commend them as a teaching approach to matrices. The ideas, methods and examples in this chapter have been written, in the main, as investigations – for the reader to be actively involved. These situations are full of extensions and possibilities and once the situation is appreciated then with active participation the reader should find much mathematical enjoyment from continuing the situation in his way with his mathematics.

7.2 Matrix description of a network

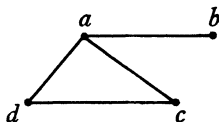


Fig. 7.1

This network can be described by the matrix

$$\begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} = \mathbf{N}$$

where the nodes a, b, c, d are written as a row and column and entries in the table (matrix) they form are a 0 if the two nodes are not joined and a 1 if they are joined. Note the 0 for node a to node a .

In this representation of a network by a matrix, the matrix is just a neat compact store of information, describing the connections within the network.

The matrix **N** can be called the 'node matrix' as it describes which nodes are directly connected.

Exercises

7.1 Describe, by a suitable matrix the network

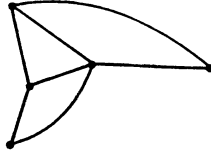


Fig. 7.2

7.2 Draw the network to describe the matrix $\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$.

Immediately, important and fundamental questions can be asked about the matrix **N** – is **N** symmetrical? Why? What do the sums of the individual rows mean?

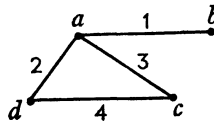


Fig. 7.3

In Fig. 7.3 the links between nodes are numbered. The 'link' matrix; the matrix describing which links are directly connected to other links is

$$\begin{matrix} & 1 & 2 & 3 & 4 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \end{matrix}$$

and the 'node-link' matrix (which need not be square) is

$$\begin{matrix} & 1 & 2 & 3 & 4 \\ \begin{matrix} a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix}$$

(The entry is 1 if the station is connected to the link, 0 otherwise.)

Can we construct a matrix to show a one-way street system? – such as for the map

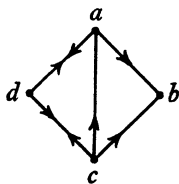


Fig. 7.4

The matrix is

$$\begin{array}{c} \text{from} \end{array} \begin{array}{c} a \\ b \\ c \\ d \end{array} \begin{array}{c} a \quad b \quad c \quad d \end{array} \begin{array}{c} \text{to} \end{array} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Note the 'from' and 'to' convention, which here needs to be specified as the matrix is now not symmetric. This convention is important and will be used in later situations.

For this matrix, what do the sums of the entries in the rows mean? What do the sums of the entries in the columns give?

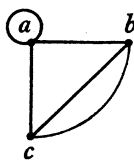


Fig. 7.5

How can a matrix show more than one link between two nodes in a network or a link around a node?

$$\begin{array}{c} a \quad b \quad c \\ a \quad b \quad c \end{array} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$$

Exercise

7.3 Draw the network described by the matrix $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix}$.

7.3 Addition of Matrices

Four towns a, b, c and d are connected as illustrated by the network

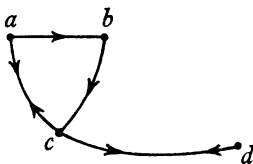


Fig. 7.6

Some new roads are planned to be added to the network. These are

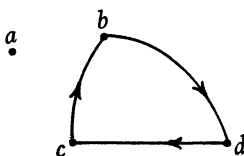


Fig. 7.7

The new network will be

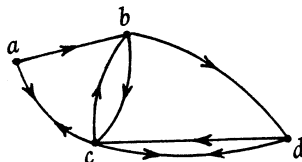


Fig. 7.8

In matrix form this situation is

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ combined with } \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ gives } \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \end{pmatrix}.$$

This method of combining two matrices is akin to addition and we write

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

where in the addition of two matrices we add corresponding entries.

The network which when added to any existing network leaves the network unchanged will be the matrix of zeros, i.e. no new roads added. Such a matrix is the identity matrix for addition.

Exercise

7.4 Add the matrices and draw the corresponding networks for

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

7.4 Multiplication of matrices

Four towns a , b , c and d are connected by bus as illustrated by the network

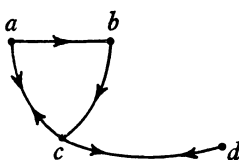


Fig. 7.9

I enjoy walking and have found walks between these towns, with the best direction for walking, as

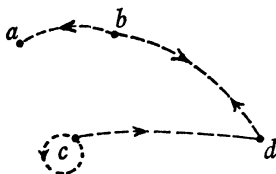


Fig. 7.10

The matrix which describes the bus connection is

$$\mathbf{B} = \begin{matrix} & \begin{matrix} a & b & c & d \end{matrix} & \text{to} \\ \begin{matrix} \text{from} \\ a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

and the matrix for the walks is

$$\mathbf{W} = \begin{matrix} & \begin{matrix} a & b & c & d \end{matrix} & \text{to} \\ \begin{matrix} \text{from} \\ a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{matrix}$$

I like going by bus from one of the towns and then taking a walk. What routes are possible?

Combining Figs. 7.9 and 7.10 we have the network as

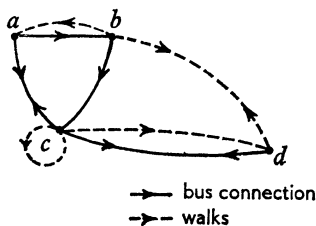


Fig. 7.11

I can take a bus to b and walk back to a from b so a to a is a possible route. a to b is not possible as I can go to c and b by bus but I have no walk from c or b to b . The matrix which shows routes between the towns when a bus journey is taken followed by a walk is, by considering the network of Fig. 7.11

$$\mathbf{M} = \begin{matrix} & \begin{matrix} a & b & c & d \end{matrix} & \text{to.} \\ \begin{matrix} \text{from} \\ a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix}$$

Can we construct the matrix $\mathbf{M}$ from the matrices $\mathbf{B}$ and $\mathbf{W}$? This situation leads to a method of combining matrices. Can we go from a to b ? asks the question – ‘is there a town x , so that one can go from a to x by bus and then from x to b by walking? We find which towns we can go to by bus *from* a by looking along the *row* a of matrix $\mathbf{B}$. We may travel to b or c . Now if we are able to walk *to* b , we look in *column* b of the matrix $\mathbf{W}$ and find we can only walk from d to b : i.e. from b or c , there is no walk to b . Hence a to b does not have a journey by bus followed by a walk, and the 0 appears in the a row, b column of the matrix $\mathbf{M}$. Consider the routes between a and d . The *row* a in $\mathbf{B}$ says we can go *from* a to b or c and the *column* d in $\mathbf{W}$ says we can come *from* b or c to d , i.e. there are two possible routes from a to d . We can find if there are routes and how many by calculation – using the 0’s and 1’s – for there is a route from a to d only when a 1 coincides in the *row* a of $\mathbf{B}$ and the *column* d of $\mathbf{W}$. The row a in $\mathbf{B}$ is 0 1 1 0 and the column d in $\mathbf{W}$ is 0

1
1
0.

The number of possible routes from a to d can be calculated from $0.0 + 1.1 + 1.1 + 0.0 = 2$, taking the corresponding entries in row a of $\mathbf{B}$ with those of column d in $\mathbf{W}$ and adding. This is *matrix multiplication*. Once the situation is appreciated, how much easier it is to obtain $\mathbf{M}$ by matrix multiplication from $\mathbf{B}$ and $\mathbf{W}$ than having to sort it out from the network! Check $\mathbf{BW} = \mathbf{M}$ by matrix multiplication. Consider $\mathbf{WB}$ by matrix multiplication and consider the result by both calculation and description (non-commutative law for matrices emerges here). Does $\mathbf{B}^2 = (\mathbf{B} \cdot \mathbf{B})$ have any meaning? ($\mathbf{B}^2$ gives the possible routes between towns such that a bus journey is followed by a bus journey.)

What happens if I take a bus journey and then a walk and then another bus journey? Will matrices and matrix multiplication cope with this situation? Try other combinations – such as $\mathbf{BWW}$. What combination of $\mathbf{B}$ and $\mathbf{W}$ will allow me to get back to my starting town? (notice principal diagonal of $\mathbf{M}$). What least combination of $\mathbf{B}$ and $\mathbf{W}$ allows communication between all the towns?

The power of the matrix application is that the rule of matrix multiplication applied in these situations gives the routes possible whatever sort of journeys are asked. There is no need to trace separate journeys; the matrix multiplication picks up the possible routes and supplies a final 'route data' matrix. Note that a bus journey ($\mathbf{B}$) followed by a walk ($\mathbf{W}$) is given by $\mathbf{BW}$ (and not $\mathbf{WB}$ – the reverse order met in transformations). This (true) order is due to our convention of writing 'from' and 'to': reverse 'from', 'to' and the conventional backward multiplication is required.

With the bus journey matrix of Fig. 7.9 what would be the walks such that a bus journey followed by a walk gave the same connections as that of the bus journeys? This walk matrix would be the identity matrix for multiplication. The bus and walk connection would need to be

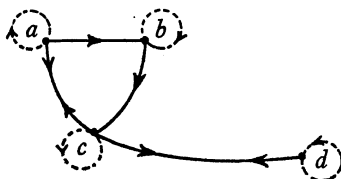


Fig. 7.12

i.e. walks around the towns themselves. The identity matrix is thus

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

– a matrix with 1's along the principal diagonals and 0's elsewhere.

7.5 When is matrix multiplication possible?

To find how the 'shapes' of two matrices must be related in order to multiply them, consider the following situation.

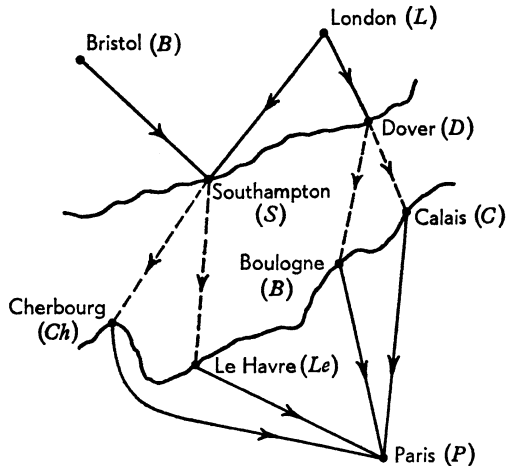


Fig. 7.13

I live in Bristol and friends live in London. We have agreed to meet in Paris, and I shall travel to Southampton and can take the channel steamer to Le Havre or Cherbourg. My friends will go to either Dover or Southampton. From Dover the channel steamer goes to Calais or Boulogne.

The matrix (M_1) giving routes from B and L to the Channel ports is

from $\begin{matrix} & S & D & to \\ B & \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \end{matrix}$ and the matrix (M_2) giving routes across the channel

is $\begin{matrix} & Le & Ch & B & C \\ S & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix}$. From the French channel ports to Paris the matrix M_3 is

$$\begin{matrix} & P \\ Le & \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\ Ch & \\ B & \\ C & \end{matrix}$$

The matrix, found from the network giving possible routes from B and L to the French ports is

$$\mathbf{M}_4 = \begin{matrix} & \begin{matrix} Le & Ch & B & C \end{matrix} \\ \begin{matrix} B \\ L \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \end{matrix}.$$

Again $\mathbf{M}_4$ can be obtained from $\mathbf{M}_1$ and $\mathbf{M}_2$ by matrix multiplication and

$$\begin{matrix} & \begin{matrix} S & D \end{matrix} \\ \begin{matrix} B \\ L \end{matrix} & \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \end{matrix} \begin{matrix} & \begin{matrix} Le & Ch & B & C \end{matrix} \\ \begin{matrix} S \\ D \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix} = \begin{matrix} & \begin{matrix} Le & Ch & B & C \end{matrix} \\ \begin{matrix} B \\ L \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \end{matrix}.$$

Note the position of S and D and how the matrices are combined. The network is simple enough to allow the possible routes from B and L to Paris to be traced and the number of routes is given by the matrix

$$(\mathbf{M}_5) = \begin{matrix} & P \\ \begin{matrix} B \\ L \end{matrix} & \begin{pmatrix} 2 \\ 4 \end{pmatrix} \end{matrix}.$$

This matrix is obtained by using the matrices $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ as

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

The matrix which describes the routes *from* Southampton or Dover *to* Bristol or London is

$$\begin{matrix} & \begin{matrix} B & L \end{matrix} & \text{to} \\ \text{from} & \begin{matrix} S \\ D \end{matrix} & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{matrix}$$

This is the matrix $\mathbf{M}_1$ with the rows of $\mathbf{M}_1$ written as columns.

The matrix formed from another matrix by writing the rows as columns is known as the *transpose* of the original matrix. The transpose of $\mathbf{M}_1$ is denoted by $\mathbf{M}_1'$. By transposing the matrices $\mathbf{M}_2$ and $\mathbf{M}_4$, show

$$\mathbf{M}_4' = (\mathbf{M}_1 \mathbf{M}_2)' = \mathbf{M}_2' \mathbf{M}_1'$$

(note the reverse order of $\mathbf{M}_1'$ and $\mathbf{M}_2'$). In general for any two matrices $\mathbf{A}$ and $\mathbf{B}$, $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.

7.6 Another situation where matrix multiplication is applicable

The 'family tree' of the male side of 'father' and 'brother' is as shown

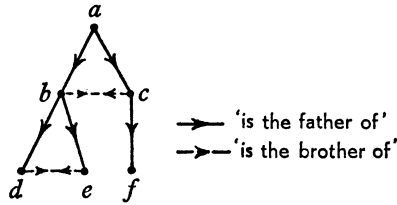


Fig. 7.14

Construct the 'father' and 'brother' matrices.

F	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>father</i>	B	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>
<i>son</i>	<i>a</i>	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$						<i>a</i>	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$					
	<i>b</i>													
	<i>c</i>													
	<i>d</i>													
	<i>e</i>													
	<i>f</i>													

Why is the matrix **B** symmetrical and **F** not symmetrical? Construct the matrix for 'uncle'.

U	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>uncle</i>
<i>nephew</i>	<i>a</i>	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$					
	<i>b</i>						
	<i>c</i>						
	<i>d</i>						
	<i>e</i>						
	<i>f</i>						

How can we construct the matrix **U** given the matrices **F** and **B**? ($\mathbf{U} = \mathbf{FB}$ – note how this depends on the ways in which **F** and **U** were written down. If **U'** had been put down, then we need a different rule than we have devised for multiplying matrices; a row on row operation if **FB** is to equal **U'**. **BF** would not even give **U'**, since we have from $\mathbf{U} = \mathbf{FB}$, then $\mathbf{U}' = (\mathbf{FB})' = \mathbf{B}'\mathbf{F}' = \mathbf{BF}'$ (as **B** is symmetrical). So if **F'** had been written down, the connection is $\mathbf{U}' = \mathbf{BF}'$ with the accepted rule for matrix multiplication. What rule would be needed to connect **F'**, **B** and **U**?).

(This example shows that there is not necessarily one rule for matrix multiplication. The rule 'row on column' is considered the easiest and is the usual one to turn up.)

7.7 Distances around networks

A salesman needs to travel through all the towns on the map. What route should be taken to travel the least distance?

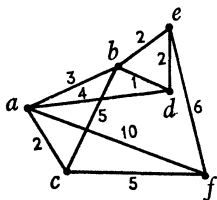


Fig. 7.15

The matrix

$$\mathbf{A} = \begin{matrix} & \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix} & \begin{pmatrix} 0 & 3 & 2 & 4 & - & 10 \\ 3 & 0 & 5 & 1 & 2 & - \\ 2 & 5 & 0 & - & - & 5 \\ 4 & 1 & - & 0 & 2 & - \\ - & 2 & - & 2 & 0 & 6 \\ 10 & - & 5 & - & 6 & 0 \end{pmatrix} \end{matrix} \quad \text{represents the distances between two connected towns.}$$

Make up a table which gives the *shortest* distance between any pair of towns via a third town.

$$\mathbf{M} = \begin{matrix} & \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix} & \begin{pmatrix} 0 & 5 & 8 & 4 & 5 & 7 \\ 5 & 0 & 5 & 4 & 3 & 8 \\ 8 & 5 & 0 & 6 & 7 & 12 \\ 4 & 4 & 6 & 0 & 3 & 8 \\ 5 & 3 & 7 & 3 & 0 & - \\ 7 & 8 & 12 & 8 & - & 0 \end{pmatrix} \end{matrix} \quad \begin{array}{l} \text{shortest distance between } a \text{ and } b \\ \text{is via } d (=5). \end{array}$$

How are the matrices $\mathbf{A}$ and $\mathbf{M}$ related? To find the shortest distance from a to b (via some other town) we take the distance *from* a to the other towns excepting b (top row of matrix $\mathbf{A}$) with the distance *to* b from the other towns (second column of matrix $\mathbf{A}$) and add the corresponding distances – the smallest will be the shortest distance from a to b via another town. Top row of $\mathbf{A}$ is 0 3 2 4 – 10 and second column of $\mathbf{A}$ is 3

0

5

1

2

–.

The corresponding additions are 3, 3, 7, 5, -, -, - so accepting the initial two 3's which represent the distance a to b , the shortest distance is 5 (via d), the other distance 7 is via c and the dashes show it is impossible to get between a and b via other towns. M can then be built up in this way. We can define M as being $A * A$ where the operation $*$ implies taking the smallest value of the corresponding additions of two elements in appropriate row and column entries of A .

What does $A * A * A$ mean?

Can the shortest route (if any!) between all the towns be found in this way? This situation is similar to critical path analysis.

Is there any situation in which taking the maximum (rather than minimum) of the corresponding additions could be relevant?

7.8 Incidence matrices

Figure 7.16 shows four stations a, b, c, d connected by 5 lines labelled 1, 2, 3, 4 and 5. The 'station-link' matrix for this network is

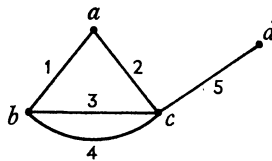


Fig. 7.16

$$A = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \begin{matrix} \text{(Station-link implies stations as rows and} \\ \text{links as columns.)} \end{matrix}$$

This matrix is an *incidence* matrix for it is made up by considering where routes and points are incident.

Exercises

- 7.5 (a) What do the sums of the individual rows of A signify?
 (b) What do the sums of the individual columns mean?
- 7.6 How can the closed loops in the network be found from the matrix A ?
- 7.7 (a) What links need to be cut to isolate the station a ?
 (b) What links need to be cut to isolate a and b (together) from the network? How can these links be found from the matrix?

- 7.8 From the transpose of $\mathbf{A}$, $\mathbf{A}'$. Compute the matrix $\mathbf{AA}'$. What information is given by the entries of $\mathbf{AA}'$?
- 7.9 For the network Fig. 7.17, label the regions α, β, γ ; γ being the whole outside region. Find the 'link-region' and 'station-region' matrices, noting them $\mathbf{P}$ and $\mathbf{Q}$ respectively. Compute and interpret $\mathbf{PP}'$, $\mathbf{P}'\mathbf{P}$, $\mathbf{QQ}'$.

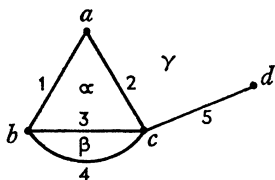


Fig. 7.17

- 7.10 Given a network we can always describe it by a matrix, but given the matrix – can we draw the corresponding network?

Draw the networks corresponding to the matrices

(i)

	1	2	3	4
a	1	0	1	1
b	1	1	0	0
c	0	1	1	0
d	0	0	0	1

(ii)

	1	2	3	4
a	1	1	0	0
b	0	1	1	0
c	1	0	1	0
d	1	0	0	1

What conditions are necessary on the matrix for a network to be drawn?

7.9 Dominance Matrices

In the House Championship Football Cup matches of four Houses A, B, C and D the results were

A had beaten B and D
 B had beaten C
 C had beaten A and D
and D had beaten B .

These results can be graphed as

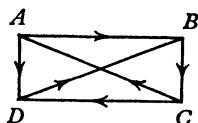


Fig. 7.18

where $A \rightarrow B$ implies ' A beats B ' i.e. A is dominant over B . We can also record the results in a matrix as

$$\mathbf{R} = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{matrix}$$

where the entry in row i , column j (a_{ij}) is a 1 if team i beats team j and a 0 if team i is beaten by team j . Since a team does not beat itself the entries in the leading diagonal will be zeros (i.e. $a_{ii}=0$). How can we find from the matrix how many games a team has won?

In the results A has beaten B and B has beaten C – so we would expect A to beat C . But this is not so – an upset of form has occurred! Therefore the relationship is not transitive (if $a_{ij}=1$ and $a_{jk}=1$ then a_{ik} may or may not be 1 – C has beaten A and A has beaten D so C should beat D – and this time has!)

If A beats B , then A is said to be *directly superior* to B . If A beats B and B beats C then A is said to be *indirectly superior* to C (even if C does beat A).

For the matches the 'indirectly superior' matrix $\mathbf{M}$ for the houses (from the network) is

$$\mathbf{M} = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix} \quad \text{Note. } C \text{ is indirectly superior to } B \text{ twice (via } D \text{ and } A).$$

Is it possible to find the matrix $\mathbf{M}$ from the matrix $\mathbf{R}$? In *row A* of $\mathbf{R}$, A has beaten (is directly superior to) B and D , and in *column C* we find C has been *beaten* by B so A is indirectly superior to C . Also in *column B*, B has been beaten by A and C so via D , A is indirectly superior to B . Thus by taking the rows of $\mathbf{R}$ with the columns of $\mathbf{R}$ and performing matrix multiplication we can compute the indirectly superior matrix, i.e. $\mathbf{M} = \mathbf{R}^2$.

The cup is won on points. If a point was given for each win we can find the number of points gained by each team by summing the rows of $\mathbf{R}$. From $\mathbf{R}$ the points results are

$$\begin{matrix} A & 2 \\ B & 1 \\ C & 2 \\ D & 1. \end{matrix}$$

Who is to have the cup? If we were to consider 'indirectly superior' also as a criterion for points, then $\mathbf{M}(=\mathbf{R}^2)$

A	2
B	2
C	3
D	1.

Putting these results together we have a points result as

A	4
B	3
C	5
D	2. So C has the cup!

The sum of the entries in a team's row of $\mathbf{R}$ and $\mathbf{R}^2$ is known as the *power* of the team, i.e. $P(A)$ is the sum of the entries in row A of $\mathbf{R} + \mathbf{R}^2$.

Consider matches between 5 teams resulting in the following directed graph.

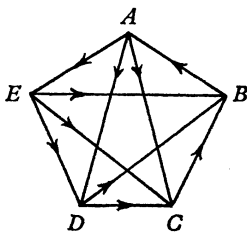


Fig. 7.19

Then

$$\mathbf{R} = \begin{array}{c} \begin{array}{ccccc} & A & B & C & D & E \end{array} \\ \begin{array}{l} A \\ B \\ C \\ D \\ E \end{array} \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix} \end{array} \quad \begin{array}{c} \text{Pts.} \\ 3 \\ 1 \\ 1 \\ 2 \\ 3 \end{array}$$

$\mathbf{R}$ gives direct superiority or *one-stage dominance*, $\mathbf{R}^2$ gives indirect superiority or *two-stage dominance*.

$$\mathbf{R}^2 = \begin{pmatrix} 0 & 3 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \end{pmatrix} \quad \begin{array}{c} \text{Pts.} \\ 6 \\ 3 \\ 1 \\ 2 \\ 4 \end{array} \quad \begin{array}{c} \text{Powers } (\mathbf{R} + \mathbf{R}^2) \\ A \\ B \\ C \\ D \\ E \end{array} \quad \begin{array}{c} 9 \\ 4 \\ 2 \\ 4 \\ 7 \end{array}$$

Note we have a top team (A) and a bottom team (C) from the 2-stage dominance.

It is possible to continue this method to find which is the superior of B and D ? Has $\mathbf{R}^3$ any meaning? $\mathbf{R}^3$ will give 3-stage dominance.

$$\mathbf{R}^3 = \begin{pmatrix} 3 & 3 & 1 & 0 & 0 \\ 0 & 3 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

But $\mathbf{R}^3$ says that A is dominant over itself 3 times (via E and B , D and B , C and B) and the other teams are also dominant over themselves as given by the leading diagonal. As this dominance is meaningless the values in the leading diagonal of $\mathbf{R}^3$ will not be counted. $\mathbf{R}^3$ becomes

	Pts.	Powers ($\mathbf{R} + \mathbf{R}^2 + \mathbf{R}^3$)
$\mathbf{R}_0^3 = \begin{pmatrix} 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 2 & 1 & 1 & 1 & 0 \end{pmatrix}$	4	13
	3	7
	2	4
	3	7
	5	12

We have still not differentiated between teams B and D . If further stages are considered more routing through the team itself becomes possible. The matrix does not show the teams used, only the number of routes between teams and if we do not count for dominance those routes through the team itself, we have to find these routes from the diagram. Although the matrix does give the number of routes to be found, the matrix method is not practical beyond the third stage.

Exercise

- 7.11** Intelligence finds that the communication links between five stations is as described in the network, Fig. 7.20. Find the matrix to describe the network. If communication is to be destroyed, which is the most vital station to destroy? Make sure of your answer by drawing the network having eliminated your station.

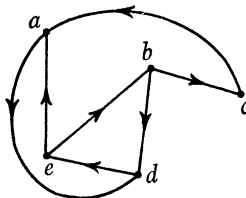


Fig. 7.20

7.10 Communication Networks

Communication between four people (*each* connected at least one-way) is described by the graph,

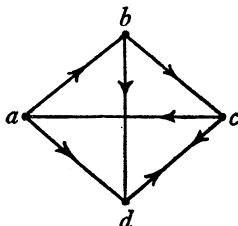


Fig. 7.21

The matrix which describes this situation is

$$\mathbf{A} = \begin{matrix} & \begin{matrix} a & b & c & d \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

where the entry a_{ij} is a 1 if i can communicate with j and a 0 otherwise, i.e. $a_{12}=1$ as a can communicate with b whereas $a_{21}=0$. Note $a_{ii}=0$; i.e. a person cannot communicate with himself! The matrix $\mathbf{A}$ describes one-stage communication between two people – and $\mathbf{A}^2$ gives 2-stage communication.

From this example (and others) it can be seen that there is always one individual who can communicate with every other individual in 1 or 2 stages. The situation in which this happens must be such that the entries exhibit the properties,

- (i) $a_{ii}=0$
- (ii) either $a_{ij}=1$ or $a_{ji}=1$ (there must be *one* communication link between any two persons). To prove that at least one individual can communicate with every other individual in 1 or 2 stages it is necessary and sufficient to show that in $\mathbf{A} + \mathbf{A}^2$ there exists one row r_i with all elements different from zero (except the element r_{ii}).

Consider the situation in which two individuals i and j *cannot* communicate in 1 or 2 stages. Then $a_{ij}=0$ in $\mathbf{A}$ and also $c_{ij}=0=\mathbf{A}_i\mathbf{A}^j$ where $\mathbf{A}_i$ is the row i in $\mathbf{A}$ and $\mathbf{A}^j$ is the column j in $\mathbf{A}$ and thus the product of the row and column (for $\mathbf{A}^2$) gives zero. Now $a_{ij}=0$ implies $a_{ji}=1$ (from initial conditions). As $\mathbf{A}_i\mathbf{A}^j=0$ then if the element $a_{ik}=1$

(the k th element in row i) then we must have the element $a_{kj}=0$ (the j th element in row k) so that the product on matrix multiplication is 0. If $a_{kj}=0$ then $a_{jk}=1$; i.e. $a_{ik}=1$ implies $a_{jk}=1$. Hence we have the situation that in a matrix $\mathbf{A}$ with n persons, the elements are such that

$$a_{j1} + a_{j2} + a_{j3} + \dots + a_{jn} \geq a_{i1} + a_{i2} + a_{i3} + \dots + a_{in} + 1$$

since we have shown if $a_{ij}=0$ then $a_{ji}=1$ and if $a_{ik}=1$ then $a_{jk}=1$ and also one of a_{ij} is considered zero (hence the 1 on the end) i.e. if one individual *cannot* communicate with another individual in 1 or 2 stages then in the matrix $\mathbf{A}$

$$a_{j1} + a_{j2} + a_{j3} + \dots + a_{jn} > a_{i1} + a_{i2} + a_{i3} + \dots + a_{in} \quad (1).$$

Now let m be the individual for whom the sum of the elements in the m th row of $\mathbf{A} + \mathbf{A}^2$ has the greatest value of all the rows. Then m can communicate with all individuals in 1 or 2 stages, for if he cannot (1) above assures us there would exist another row with a greater sum (since we began with the premise that one individual *could not* communicate with another individual in 1 or 2 stages and were led to result (1)) and this contradicts the statement that the row m has the greatest sum.

Exercises

7.12 if

$$\mathbf{A} = \begin{matrix} & \begin{matrix} a & b & c & d & e \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \end{matrix}$$

- (i) draw the directed network
- (ii) calculate $\mathbf{A}^2$
- (iii) is the network such that $a_{ii}=0$ and either $a_{ij}=1$ or $a_{ji}=1$
- (iv) which individual can communicate with all others in 1 or 2 stages, as selected by the 'greatest sum' (check result on graph)
- (v) which other individuals (if any) can communicate with all others in 1 or 2 stages.

7.13 *Rumours*: Mrs. A chatters to Mrs. B , Mrs. C and Mrs. F . Mrs. C chatters to Mrs. A , Mrs. E and Mrs. D . Mrs. E chatters to Mrs. C and Mrs. F . Mrs. F chatters to Mrs. A and Mrs. E , and Mrs. D chatters to Mrs. C and Mrs. E . Draw a 'chatter diagram' for these communications. Construct a matrix showing rumours at first and second hand and decide who would spread a rumour the quickest throughout the group.

7.11 Sociometric Data

The social attitude between individuals may be summarised into three categories – acceptance, indifference or rejection. From a study of the social habits of individuals it is possible to develop a sociogram – a diagram which indicates the feelings of one individual towards another. In such a diagram:

$a \rightarrow b$; a indicates acceptance of b – true friendship

$a - b$; a indicated indifference towards b

$a \dashv b$; a indicates rejection of b – not on talking terms.

A sociogram between 3 individuals may be as in

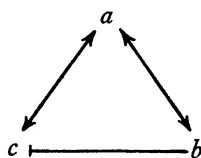


Fig. 7.22

a accepts b and c

b accepts a and rejects c

c accepts a and indifferent to b .

By means of matrices it is possible to find a method of quantifying such a sociogram to show

- (i) a number status for each individual in the group taking into account *his* choices and rejections and also the kinds of choices of others in which he is involved.
- (ii) ranking an individual showing his place in the group giving an indication of social adjustment or social rejection
- (iii) determination of cliques within a group
- (iv) establishment of a numerical index for the group so it may be compared to other groups, to an optimum value for a 'perfect' group and to compare the same group with itself at a later stage when different sociometric information is collected.

The matrices which describe the sociogram Fig. 7.22 are

$$\mathbf{A} = \text{acceptance} \quad \begin{matrix} & a & b & c \\ \begin{matrix} a \\ b \\ c \end{matrix} & \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \end{matrix}$$

where a 1 indicates acceptance and 0 indifference. The *rows* show what

choices a person has made and the *columns* the persons who have chosen the head of the column.

With **A** is the rejection matrix

$$\mathbf{R} = \text{rejection} \begin{array}{c} \begin{array}{ccc} & a & b & c \\ a & 0 & 0 & 0 \\ b & 0 & 0 & -1 \\ c & 0 & 0 & 0 \end{array} \end{array}$$

where a 0 indicates indifference and -1 rejection. Row and column meaning as before.

Consider

$$\mathbf{A}^2 = \begin{array}{c} \begin{array}{ccc} & a & b & c \\ a & 2 & 0 & 0 \\ b & 0 & 1 & 1 \\ c & 0 & 1 & 1 \end{array} \end{array}$$

which gives 'two-stage' acceptance, i.e. *b* is acceptable to *c* via *a*.

A² describes relations within a group via another person in the group – even showing when a person is *mutually* acceptable e.g. *a* (twice) and *b* and *c* (once). These are the values in the leading diagonal of **A**² and are the clues to any cliques.

Now

$$\mathbf{A} + \mathbf{A}^2 = \begin{array}{c} \begin{array}{ccc} & a & b & c \\ a & 2 & 1 & 1 \\ b & 1 & 1 & 1 \\ c & 1 & 1 & 1 \end{array} \end{array}$$

gives the 'one'- and 'two'-stage relationships between this group of individuals.

The total of the *columns* give the number of 'one'- and 'two'-stage *inward* relationships (i.e. the individual has been chosen by these number of people) enjoyed by each individual – *a*, 4 and *b* and *c*, 3. The *row* sums give the *outward* relationships between the individuals: *a*, 4 and *b* and *c*, 3.

If we now use **R** and compute

$$\mathbf{A} + \mathbf{A}^2 + \mathbf{R} = \begin{array}{c} \begin{array}{ccc} & a & b & c \\ a & 2 & 1 & 1 \\ b & 1 & 1 & 0 \\ c & 1 & 1 & 1 \end{array} \end{array}$$

we have a matrix from which social values can be determined – the column totals giving number of inward relationships and the rows totals outward relationships. A social index can be found for each individual

by combining his row and column total – and if it is considered that outward relationships are more important than inward relationships we can multiply the row totals by two – before adding to the column total. Here then the social index of an individual is the addition of his column total with twice his row total in the matrix $\mathbf{A} + \mathbf{A}^2 + \mathbf{R}$. In the example

$$a = 4 + 2(4) = 10$$

$$b = 3 + 2(2) = 7$$

$$c = 2 + 2(3) = 8$$

and these are the social indices for each individual. The composite social index for the group is $10 + 7 + 8 = 25$. Ideally when each individual accepts all others, the group index is $3n^2(n-1)$ when n is the number of individuals in the group – i.e. for a threesome 54.

Cliques can be discovered from mutual choices. If we find the individuals who have the greatest social index and from $\mathbf{A}^2$ the individuals who enjoy most mutual choice, then from $\mathbf{A}$ we can decide who those individuals are and decide if they form a clique.

Exercise

7.14 Investigate the sociogram

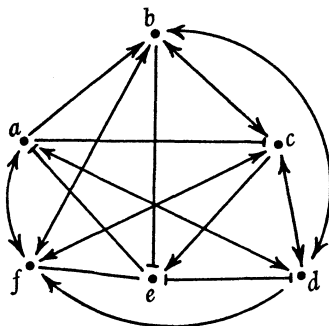


Fig. 7.23

(The methods given in 7.11 were suggested by an article by F. Schippert in the American Journal, *School Science and Mathematics*, Dec. 1966.)

8

SUCCESSIVE EVENTS, STOCHASTIC PROCESSES AND MARKOV CHAINS

8.1 The Use of Matrices in Life cycles

The following example, although somewhat artificial, shows how matrices can describe situations in which events recur. Other more realistic situations will be dealt with later, but this is a good situation with which to begin.

A certain species of animal has a life cycle such that $\frac{2}{3}$ of the animals in their 1st year of life survive their first birthday; of these, $\frac{1}{2}$ survive their second birthday and live into their third year; no animal lives more than three years.

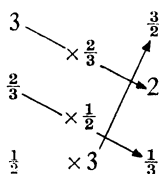
During the third year of life, reproduction of the species takes place, and during this year three animals (on the average) are born for each animal alive. If we start with 3,000 animals, with 1,000 in each year group, how will the animal population fluctuate over future years?

The starting situation working in thousands and writing as a column vector is $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ for the 1st, 2nd and 3rd years respectively. After the first year, the thousand animals in their first year will become $\frac{2}{3} \cdot 1,000$ into the 2nd year (since only $\frac{2}{3}$ of the animals survive their first birthday). The thousand animals in the 2nd year become $\frac{1}{2} \cdot 1,000$ in the third year and the 1,000 animals in the third year reproduce $3 \times 1,000$ to go into the 1st year and no animal lives into the 4th year and beyond.

The situation now is $\begin{pmatrix} 3 \\ \frac{2}{3} \\ \frac{1}{2} \end{pmatrix}$ for the 1st, 2nd and 3rd years.

After the second year in the life cycle, the situation is $\begin{pmatrix} \frac{3}{2} \\ 2 \\ \frac{1}{3} \end{pmatrix}$.

These values are found from the first year values by a calculation explained below, which can be shown schematically in the diagram.



We have taken the first year value, multiplied it by $\frac{2}{3}$ and placed it in the second year, the 2nd year value is multiplied by $\frac{1}{2}$ and put in the 3rd year, and the 3rd year value is multiplied by 3 and put into the 1st year. A matrix has the property of multiplying values and transposing the places when multiplied with another matrix and it is therefore possible to represent the life cycle of this species of animal by a matrix.

Consider $\begin{pmatrix} 0 & 0 & c \\ a & 0 & 0 \\ 0 & b & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} cz \\ ax \\ by \end{pmatrix}$ - note the transposition of x, y

and z and the positions of a, b and c in the matrix. If the matrix $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$

is taken as the values after the first year $\begin{pmatrix} 3 \\ \frac{2}{3} \\ \frac{1}{2} \end{pmatrix}$ and the matrix $\begin{pmatrix} 0 & 0 & c \\ a & 0 & 0 \\ 0 & b & 0 \end{pmatrix}$

as $\begin{pmatrix} 0 & 0 & 3 \\ \frac{2}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$ with the proportions surviving in the year groups, then

we can find the number of animals in the year groups of the second year by

$$\begin{pmatrix} 0 & 0 & 3 \\ \frac{2}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 3 \\ \frac{2}{3} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ 2 \\ \frac{1}{3} \end{pmatrix}.$$

The matrix $\mathbf{M} = \begin{pmatrix} 0 & 0 & 3 \\ \frac{2}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$ completely describes the life cycle of the animals.

After three years the situation is given by $\mathbf{M} \begin{pmatrix} \frac{3}{2} \\ 2 \\ \frac{1}{3} \end{pmatrix}$ that is

$$\begin{pmatrix} 0 & 0 & 3 \\ \frac{2}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{2} \\ 2 \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

We have arrived at the starting situation and the cycle then repeats. Show $\mathbf{M}^3 = \mathbf{I}$.

Examples

8.1 Find the total number of animals in each of the three years of the above cycle.

8.2 Investigate and describe the life cycle represented by the matrices.

$$(a) \begin{pmatrix} 0 & 0 & 6 \\ \frac{2}{3} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \end{pmatrix} \quad (b) \begin{pmatrix} 0 & 3 \\ \frac{1}{3} & 0 \end{pmatrix} \quad (c) \begin{pmatrix} 0 & 0 & 0 & 24 \\ \frac{1}{3} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \end{pmatrix}$$

8.2 Survival or Extinction of the Species

A life cycle is represented by the matrix $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix}$; that is the species is such that $\frac{1}{4}$ survive their 1st birthday and live into their second year and for each individual alive in the second year 1 is born and none of the species lives longer than 2 years. How does the population fluctuate over future years?

Assume that 100 of the species is alive in the 1st and 2nd years as a starting situation.

	Total
After 1 year the position is $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{4} \end{pmatrix}$	$1\frac{1}{4}$
After 2 years the position is $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{16} \end{pmatrix}$	$\frac{1}{2}$
After 3 years the position is $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{4} \\ \frac{1}{16} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{64} \end{pmatrix}$	$\frac{5}{16}$
After 4 years the position is $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{4} \\ \frac{1}{64} \end{pmatrix} = \begin{pmatrix} \frac{1}{16} \\ \frac{1}{64} \end{pmatrix}$	$\frac{5}{64}$

Will the cycle repeat?

After 2 years the position can be seen as

$$\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{1}{4} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix}^2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

After 3 years the position is $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix}^3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$

After n years the position is $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix}^n \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$

From the pattern of the above results the total population of the species as given by $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & 0 \end{pmatrix}^n \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ can be seen to be decreasing as n increases.

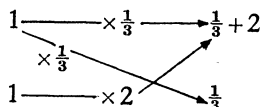
Example

8.3 Investigate the situation as described by the matrix $\begin{pmatrix} 0 & 6 \\ \frac{2}{3} & 0 \end{pmatrix}$. Can you conjecture as to the future of a species in the general case as described by the matrix $\begin{pmatrix} 0 & b \\ a & 0 \end{pmatrix}$ for various values of a and b ?

8.3 'Steady State' Populations

We investigate the yearly population of a species as described by the matrix $\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix}$. The species is such that for every three individuals in their first year one individual is born, (represented by the $\frac{1}{3}$ in the matrix) and $\frac{1}{3}$ of those in their first year survive their 1st birthday and live into the second year. During the second year of life each individual bears two offspring, and no individual survives the second year.

Assuming 100 individuals in each of the 1st and 2nd years as a starting situation, after 1 year the position is



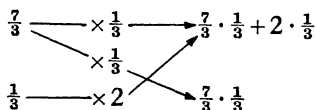
since for every 3 individuals in the first year 1 is born to go into the next first year and $\frac{1}{3}$ survive into the second year of life, and of those in their second year, 2 individuals are born for each in their second year and none survive into the third year.

The matrix gives the first year situation as

$$\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} + 2 \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{7}{3} \\ \frac{1}{3} \end{pmatrix}$$

corresponding to $2\frac{2}{3}$ (hundreds).

For the second year the situation is



Or from the matrix

$$\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} \frac{7}{3} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \cdot \frac{7}{3} + 2 \cdot \frac{1}{3} \\ \frac{1}{3} \cdot \frac{7}{3} \end{pmatrix} = \begin{pmatrix} \frac{13}{9} \\ \frac{7}{9} \end{pmatrix} \quad \text{total } 2\frac{2}{9}.$$

For the third year the position is

$$\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} \frac{13}{9} \\ \frac{7}{9} \end{pmatrix} = \begin{pmatrix} \frac{55}{27} \\ \frac{13}{27} \end{pmatrix} \quad \text{total } 2\frac{14}{27}.$$

And for the fourth year the position is

$$\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} \frac{55}{27} \\ \frac{13}{27} \end{pmatrix} = \begin{pmatrix} \frac{133}{81} \\ \frac{55}{81} \end{pmatrix} \quad \text{total } 2\frac{26}{81}.$$

Collected together the number of each of the species in each year of its two year cycle and the population in each successive year is

	1st year of life	2nd year of life	total in year (100's)
Start	1	1	2
1	$\frac{7}{3} = 2.33$	$\frac{1}{3} = 0.33$	$2\frac{2}{3} = 2.67$
2	$1\frac{4}{9} = 1.44$	$\frac{7}{9} = 0.78$	$2\frac{2}{9} = 2.22$
3	$\frac{55}{27} = 1.94$	$\frac{13}{27} = 0.48$	$2\frac{14}{27} = 2.51$
4	$\frac{133}{81} = 1.64$	$\frac{55}{81} = 0.68$	$2\frac{26}{81} = 2.41$
			etc.

There is much fluctuation in the values here. It appears that the population is neither decreasing or increasing and may be tending towards a limiting value – ‘a steady state’ value. Can we find this value?

If the population of the species does reach a ‘steady-state’ after some period of time, then the situation would be that if the numbers in each year at this stage were $\begin{pmatrix} x \\ y \end{pmatrix}$ then $\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$ that is, the ‘transition’ matrix $\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix}$ describing the change in population from year to year has not altered the value of the vector $\begin{pmatrix} x \\ y \end{pmatrix}$ – the ‘steady state’.

This vector must be an *eigenvector* of the matrix. In this situation we see the importance of the eigenvector; it gives those values of the situation, wherein after the matrix is continually applied, the matrix does not alter the eigenvector values.

If T is the matrix $\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix}$ and $\mathbf{x}$ the vector $\begin{pmatrix} x \\ y \end{pmatrix}$ then the steady-state position is $T\mathbf{x} = \mathbf{x}$ where $\mathbf{x}$ is the eigenvector of the matrix T .

If λ is an eigenvalue of the matrix $\mathbf{T} = \begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix}$ then

$$\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}.$$

So

$$\frac{1}{3}x + 2y = \lambda x$$

and

$$\frac{1}{3}x + 0 \cdot y = \lambda y$$

from which

$$\begin{cases} (\frac{1}{3} - \lambda)x + 2y = 0 \\ \frac{1}{3}x + (0 - \lambda)y = 0 \end{cases} \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (1).$$

For these equations to have non-zero solutions we must have

$$\begin{pmatrix} \frac{1}{3} - \lambda & 2 \\ \frac{1}{3} & 0 - \lambda \end{pmatrix} = 0$$

that is $-\lambda(\frac{1}{3} - \lambda) - \frac{2}{3} = 0$; $\lambda^2 - \frac{1}{3}\lambda - \frac{2}{3} = 0$ and $\lambda = 1$ or $-\frac{2}{3}$.

From (1), when $\lambda = 1$ the eigenvector is given by $(\frac{1}{3} - 1)x + 2y = 0$, i.e. $\frac{2}{3}x = 2y$ or $x = 3y$ and as a vector by $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

Then

$$\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

But since $\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} 3k \\ k \end{pmatrix} = \begin{pmatrix} 3k \\ k \end{pmatrix}$ we can only deduce that when the steady state is reached the number of the species in the first year of life is three times the number in the second year. We are unable to find the actual values of the population in their first and second year of life when the steady state is reached.

To be able to find these values of the steady state we need to be able to compute successive powers of the matrix, which will give the population values for each successive year, and then endeavour to deduce from

these values the steady state numbers. Since $\mathbf{T} = \begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix}$ and the popu-

lation for the first year is given by $\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and the second year by

$\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix}^2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and the n th year by $\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix}^n \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ we require a method of

computing $\begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix}^n$.

8.4 Power of a Matrix

The other eigenvalue of $\mathbf{T}$ was $-\frac{2}{3}$ and the corresponding eigenvector is given by $(\frac{1}{3} + \frac{2}{3})x + 2y = 0$, that is $x = -2y$, or by the vector $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$

$$\left[\text{check } \begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{4}{3} \\ -\frac{2}{3} \end{pmatrix} = -\frac{2}{3} \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right].$$

The two eigenvectors of $\mathbf{T}$ are given by the vectors $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$.

Construct the matrix $\mathbf{P}$ by writing the eigenvectors as *columns*, that is

$$\mathbf{P} = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}.$$

Then the inverse of $\mathbf{P}$ is $\mathbf{P}^{-1} = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$ and computing $\mathbf{P}^{-1}\mathbf{TP}$ gives

$$\begin{aligned} \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} &= \frac{1}{5} \begin{pmatrix} 1 & 2 \\ \frac{2}{3} & -2 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 5 & 0 \\ 0 & -\frac{10}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{2}{3} \end{pmatrix}. \end{aligned}$$

This is a matrix of diagonal form with the eigenvalues of $\mathbf{T}$ as the entries on the major diagonal. (It will be proved later that using the eigenvectors of a matrix in the above way will always give a result of this form.)

$$\text{Let } \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{2}{3} \end{pmatrix} \text{ then } \mathbf{P}^{-1}\mathbf{TP} = \mathbf{D}.$$

We require to find $\mathbf{T}^n$.

Now $\mathbf{P}(\mathbf{P}^{-1}\mathbf{TP})\mathbf{P}^{-1} = \mathbf{PDP}^{-1}$, pre- and post-multiplying the equation $\mathbf{P}^{-1}\mathbf{TP} = \mathbf{D}$ by $\mathbf{P}$ and $\mathbf{P}^{-1}$ respectively, that is

$$(\mathbf{PP}^{-1})\mathbf{T}(\mathbf{PP}^{-1}) = \mathbf{PDP}^{-1}$$

(by the associative law) and thus

$$\mathbf{T} = \mathbf{PDP}^{-1}.$$

Then

$$\mathbf{T}^2 = (\mathbf{PDP}^{-1})(\mathbf{PDP}^{-1}) = \mathbf{PD}(\mathbf{P}^{-1}\mathbf{P})\mathbf{D}\mathbf{P}^{-1} = \mathbf{PD}^2\mathbf{P}^{-1}.$$

$$\mathbf{T}^3 = \mathbf{PD}^2\mathbf{P}^{-1}(\mathbf{PDP}^{-1}) = \mathbf{PD}^3\mathbf{P}^{-1}$$

and hence (by induction)

$$\mathbf{T}^n = \mathbf{PD}^n\mathbf{P}^{-1}$$

$$\mathbf{T}^n = \begin{pmatrix} \frac{1}{3} & 2 \\ \frac{1}{3} & 0 \end{pmatrix}^n = \mathbf{PD}^n\mathbf{P}^{-1} = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -\frac{2}{3} \end{pmatrix}^n \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}.$$

It is easy to compute powers of a matrix in diagonal form since

$$\begin{pmatrix} 1 & 0 \\ 0 & -\frac{2}{3} \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{2}{3} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & (-\frac{2}{3})^2 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 0 \\ 0 & -\frac{2}{3} \end{pmatrix}^3 = \begin{pmatrix} 1 & 0 \\ 0 & (-\frac{2}{3})^3 \end{pmatrix}$$

and by induction

$$\begin{pmatrix} 1 & 0 \\ 0 & -\frac{2}{3} \end{pmatrix}^n = \begin{pmatrix} 1 & 0 \\ 0 & (-\frac{2}{3})^n \end{pmatrix}.$$

Hence

$$\begin{aligned} \mathbf{T}^n &= \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (-\frac{2}{3})^n \end{pmatrix}^{\frac{1}{5}} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -2(-\frac{2}{3})^n \\ 1 & (-\frac{2}{3})^n \end{pmatrix}^{\frac{1}{5}} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3 + 2(-\frac{2}{3})^n & 6 - 6(-\frac{2}{3})^n \\ 1 - (-\frac{2}{3})^n & 2 + 3(-\frac{2}{3})^n \end{pmatrix}. \end{aligned}$$

Now as n (the number of years in the life of the species) increases from the starting situation $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{T}^n$ tends to the limiting state

$$\begin{aligned} &\frac{1}{5} \begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (\text{since } (-\frac{2}{3})^n \rightarrow 0 \text{ as } n \rightarrow \infty) \\ &= \frac{1}{5} \begin{pmatrix} 9 \\ 3 \end{pmatrix} = \begin{pmatrix} \frac{9}{5} \\ \frac{3}{5} \end{pmatrix}. \end{aligned}$$

Thus the population will tend towards a limiting value of $\frac{9}{5} + \frac{3}{5} = 2.4$ (hundreds), distributed in the proportions $(\frac{9}{5}, \frac{3}{5}) = (1.8, 0.6)$ in the first and second years of life.

Compare this result with the calculated values for the first four years of the life cycle, and note the values $(\frac{9}{5}, \frac{3}{5})$ agree with the eigenvector of the matrix. In this example the eigenvalue $\lambda = 1$ and a steady state exists; it is not difficult to deduce that if $\lambda > 1$ the population will be increasing and that if $\lambda < 1$ the population is decreasing.

Examples

8.4 If λ_1 and λ_2 are the eigenvalues of the characteristic equation of the matrix $\mathbf{T} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ and $\begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ the corresponding eigenvectors, then the matrix $\mathbf{P}$ constructed by writing the eigenvectors as columns is $\begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}$. Show $\mathbf{P}^{-1}\mathbf{TP} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

This establishes, as a general result, the method used in the example to reduce $\mathbf{T}$ to diagonal form.

8.5 Let $\mathbf{D} = \mathbf{P}^{-1}\mathbf{TP}$ then $\mathbf{T}^n = \mathbf{PD}^n\mathbf{P}^{-1}$

$$\begin{aligned} &= \mathbf{P} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^n \mathbf{P}^{-1} \\ &= \mathbf{P} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \mathbf{P}^{-1}. \end{aligned}$$

Insert the matrix $\mathbf{P}$ and compute the full result of $\mathbf{T}^n$. Hence deduce that if $\lambda_1 = 1$ and $\lambda_2 < 1$ a 'steady-state' is eventually reached; if $\lambda_1 > 1$ the population will continually increase and if $\lambda_1 < 1$ and $\lambda_2 < 1$ the population will continually decrease and become extinct.

8.6 In the following examples describe the life cycle as given by the matrix and then consider the future population expectancy of the species. If a 'steady-state' value is to be reached, find this value.

$$(a) \begin{pmatrix} \frac{1}{2} & 2 \\ \frac{1}{4} & 0 \end{pmatrix} \quad (b) \begin{pmatrix} \frac{1}{4} & 3 \\ \frac{1}{8} & 0 \end{pmatrix} \quad (c) \begin{pmatrix} 1 & 3 \\ \frac{1}{4} & 0 \end{pmatrix}.$$

(Investigate how these matrices are devised – the general matrix is of the form $\begin{pmatrix} a & c \\ b & 0 \end{pmatrix}$).

8.7 The theory of 3×3 matrices describing life-cycles is more complicated. The reader is invited to study matrices of the form

$$(i) \begin{pmatrix} 0 & 0 & c \\ a & 0 & 0 \\ 0 & b & 0 \end{pmatrix} \quad \text{and} \quad (ii) \begin{pmatrix} 0 & d & c \\ a & 0 & 0 \\ 0 & b & 0 \end{pmatrix};$$

e.g.

$$\begin{pmatrix} 0 & 0 & 4 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 & 6 \\ \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

These will be found to describe life cycles in which there will be only one stable distribution of *proportions* of the population in each year of its life and there is no total population which the species will finally tend towards. (Since the *only* eigenvector comes from $\lambda = 1$ and it is impossible then to diagonalise the matrix and consider powers as $n \rightarrow \infty$).

8.5 Stochastic Processes and Markov Chains

In the previous section we have seen how a matrix may describe the outcome of successive events (the reproduction and dying in the life cycle of a species). The matrix is often called the *transition matrix* as it describes the situation after each operation. A sequence of events where

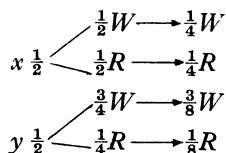
the outcome is dependent on some chance element is known as a *stochastic process*. In these events probability is involved and the entries in the transition matrix will be probabilities. We examine some examples of stochastic processes.

(i) A black bag (x) contains 2 white and 2 red balls and another black bag (y) contains 3 white and 1 red ball. If both bags are held out, what is the probability a white ball is drawn out of one bag? There are *two* events to be considered here; one, the choosing of bag x or y and second the choosing of a ball in the bag. The probability matrix for choosing a bag is $\begin{smallmatrix} x \\ y \end{smallmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ since it is assumed that each bag has an equal likelihood of

being chosen. The probability matrix for the balls is $\begin{smallmatrix} & x & y \\ W & \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \\ R & \begin{pmatrix} \frac{3}{4} \\ \frac{1}{4} \end{pmatrix} \end{smallmatrix}$ since in x the probabilities of choosing a white or a red is $\frac{1}{2}$ whilst in a bag y the probability of choosing a white is $\frac{3}{4}$ and the red, $\frac{1}{4}$. By combining these matrices we shall have the probabilities of choosing a white or red ball as

$$\begin{pmatrix} \frac{1}{2} & \frac{3}{4} \\ \frac{1}{2} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \times \frac{1}{2} + \frac{3}{4} \times \frac{1}{2} \\ \frac{1}{2} \times \frac{1}{2} + \frac{1}{4} \times \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{5}{8} \\ \frac{3}{8} \end{pmatrix}.$$

Hence the probability of choosing a white ball is 5 in 8 times. (Note it is assumed also in this situation that if the experiment is continually performed, to see how it compares in practice with the theory, the ball chosen is replaced before the next draw is made.) The matrix multiplication gives the probable outcomes since as the probability of choosing bag x is $\frac{1}{2}$ and in bag x there is a probability of $\frac{1}{2}$ of choosing a white ball, so there is a probability of $\frac{1}{4}$ in this case of getting a white ball and for bag y the probability of choosing it is $\frac{1}{2}$ together with the probability of $\frac{3}{4}$ of obtaining a white ball having chosen y , that is probability of $\frac{3}{8}$ of getting a white ball from y ; the total probability of obtaining a white ball is $\frac{1}{4} + \frac{3}{8} = \frac{5}{8}$. A 'tree diagram' shows how the probabilities can be obtained and how matrix multiplication performs the same operation.



The matrix multiplication deals with the situation since we have two *successive* events of which the transition probabilities are known.

(ii) From a vending machine I can get nut, raisin or plain chocolate bars. If I had a nut bar I am equally likely to choose a raisin or plain the next time. If I had the raisin I am twice as likely to choose the plain

as the nut for the next; and if I had the plain, next time I would always get the nut bar! What is the matrix which describes my choices?

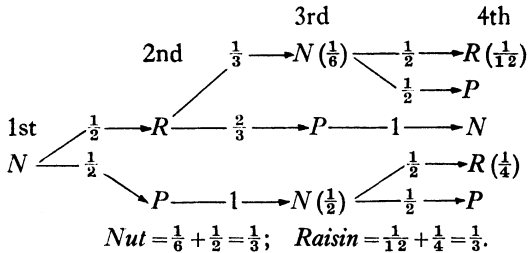
$$\begin{array}{cc} & \begin{array}{ccc} N & R & P \end{array} \\ \begin{array}{c} \text{2nd} \\ \text{choice} \end{array} & \begin{array}{ccc} N & \left(\begin{array}{ccc} 0 & \frac{1}{3} & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{2}{3} & 0 \end{array} \right) & P \end{array} \end{array} \quad \begin{array}{c} \text{1st choice} \end{array}$$

If my first choice is a nut bar, what is the probability I have a nut bar on my 3rd choice and what is the probability I have a raisin bar on my 4th choice? Note that the starting situation is given and also how the matrix multiplication deals with the probability outcomes and check it is correct by general probability methods.

$$\text{1st choice: probability vector} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{array}{c} N \\ R \\ P \end{array}$$

$$\text{2nd choice: probability vector} = \begin{pmatrix} 0 & \frac{1}{3} & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{array}{c} N \\ R \\ P \end{array}$$

$$\text{3rd choice: probability vector} = \begin{pmatrix} 0 & \frac{1}{3} & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{2}{3} & 0 \end{pmatrix}^2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{array}{c} N \\ R \\ P \end{array}$$



(iii) Today is the first day of term. A maths and physics student knows that today there is an even chance of working at maths or physics. But tomorrow will be different! If he works at maths today there is only a probability of $\frac{1}{5}$ that he will work at maths the next day (and therefore a probability of $\frac{4}{5}$ that he will switch and do physics). On the other hand, if today he works at physics there is a probability of $\frac{3}{5}$ that he will do physics the next day. These probabilities continue for subsequent days.

At the beginning of the 1st day the probability matrix is

$$\begin{array}{c} M \\ P \end{array} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

At the beginning of the 2nd day the transition matrix is

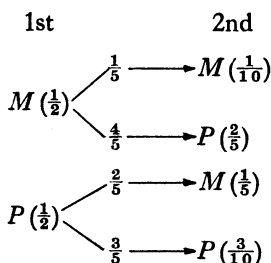
$$\begin{array}{cc} & \begin{matrix} M & P \end{matrix} \text{ today} \\ \text{tomorrow} & \begin{pmatrix} M & P \\ \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \end{array}$$

(where the entry $\frac{2}{5}$ states that if the student did physics today there is a probability of $\frac{2}{5}$ he will do maths tomorrow).

What is the probability of doing maths on the second day?

$$\begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{3}{10} \\ \frac{7}{10} \end{pmatrix}.$$

Again check the matrix multiplication deals correctly with the probabilities



What are the probabilities for the 3rd day?

$$\begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}^2 \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} \frac{3}{10} \\ \frac{7}{10} \end{pmatrix} = \begin{pmatrix} \frac{17}{50} \\ \frac{33}{50} \end{pmatrix}.$$

Questions such as (ii) and (iii) are answered by taking powers of the matrix. In example (iii) what is the probability of the student working at maths at the end of the term – after say 100 days? To multiply the matrix $\begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$ by itself 100 times we use our earlier method of computing powers of a matrix.

If $\mathbf{A} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$ then the eigenvalues of $\mathbf{A}$ are given by

$$(\frac{1}{5} - \lambda)(\frac{3}{5} - \lambda) - \frac{8}{25} = 0$$

or

$$\lambda^2 - \frac{4}{5}\lambda - \frac{1}{5} = 0$$

i.e.

$$5\lambda^2 - 4\lambda - 1 = 0$$

and

$$(5\lambda + 1)(\lambda - 1) = 0,$$

so

$$\lambda = 1 \text{ or } -\frac{1}{5}.$$

For the eigenvectors, from $(\frac{1}{5} - \lambda)x + \frac{2}{5}y = 0$ when $\lambda = 1$ the eigenvector is $2x - y = 0$, or the vector $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. When $\lambda = -\frac{1}{5}$ the eigenvector is $x + y = 0$, or the vector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Let $\mathbf{P}$ be the matrix whose columns are the eigenvectors, that is $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$. Then from work in the previous section

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{5} \end{pmatrix}$$

which is a diagonal matrix with the eigenvalues as principal diagonal entries. As before $\mathbf{A}^n = \mathbf{P}\mathbf{D}^n\mathbf{P}^{-1}$

$$\begin{aligned} &= \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (-\frac{1}{5})^n \end{pmatrix}^{\frac{1}{3}} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 1 & (-\frac{1}{5})^n \\ 2 & -(-\frac{1}{5})^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 1 + 2(-\frac{1}{5})^n & 1 - (-\frac{1}{5})^n \\ 2 - 2(-\frac{1}{5})^n & 2 + (-\frac{1}{5})^n \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{3} + \frac{2}{3}(-\frac{1}{5})^n & \frac{1}{3} - \frac{1}{3}(-\frac{1}{5})^n \\ \frac{2}{3} - \frac{2}{3}(-\frac{1}{5})^n & \frac{2}{3} + \frac{1}{3}(-\frac{1}{5})^n \end{pmatrix}. \end{aligned}$$

When $n = 100$ the entries closely approximate to $\begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{2}{3} \end{pmatrix}$ (and if the term is endless they do!).

On the last day of term the probabilities of the student working on maths or physics are given by $\begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix} \frac{M}{P}$ so there is a 2 to 1 chance he will be doing physics.

Such problems as these are known as 'Markov chains' - where the probability at a given instance is a function of the outcome immediately preceding.

Examples

8.8 It has been found that each year $\frac{2}{3}$ of the population in a town moves into the country whilst $\frac{1}{3}$ of the country dwellers move into the town. Consider the future balance of population between town and country, assuming these proportions of change continue. Consider 7,000 town dwellers and 7,000 country dwellers as a starting position.

8.9 On the 1st day of term a lecturer has a probability of $\frac{1}{3}$ of being late for his lecture. But if he is late there is a probability of $\frac{3}{4}$ that he will be early for the next lecture, and if he arrives early there is a probability of $\frac{2}{3}$ he will be early next time. If these probabilities continue for subsequent days, what is the probability of the lecturer being late at the end of term lecture?

8.10 Consider the general theory behind these examples. Each transition matrix is such that the sum of the elements in each column is 1 – why? When a transition matrix is compounded with another the sum of each column remains 1 – why?

The 2×2 matrices in the examples must be of the form $\begin{pmatrix} a & 1-b \\ 1-a & b \end{pmatrix}$

where $0 < a < 1$ and $0 < b < 1$. Show the eigenvalues of this matrix are 1 and $a+b-1$. Since an eigenvalue is always 1, and $a+b-1 < 1$, each situation will approach some stable state (it will not be possible to have a situation where one aspect is absorbed by the other).

8.6 The use of the eigenvector in stochastic processes

As with the life cycles we are looking for the stable state when the powers of the matrix no longer change the matrix, that is we have reached the state when $\mathbf{TP} = \mathbf{P}$ where $\mathbf{T}$ is the transition matrix and $\mathbf{P} = \mathbf{T}^n$. Again we are considering the eigenvectors of the matrix; for the matrix $\mathbf{P}$ must have entries of the eigenvectors of $\mathbf{T}$ so that $\mathbf{TP} = \mathbf{P}$. Finding the eigenvectors of $\mathbf{T}$ will supply the steady-state proportions and there will be no need to compute powers of the matrix.

Consider the worked example in the text of the student studying maths and physics. The transition matrix $\mathbf{T}$ is $\begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$. Let an eigenvector of $\mathbf{T}$ be $\begin{pmatrix} x \\ y \end{pmatrix}$. Then we require the situation

$$\begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{when } \lambda = 1.$$

Then

$$\frac{1}{5}x + \frac{2}{5}y = x \quad \text{and} \quad \frac{4}{5}x + \frac{3}{5}y = y$$

so that

$$-\frac{4}{5}x + \frac{2}{5}y = 0 \quad \text{and} \quad \frac{4}{5}x - \frac{2}{5}y = 0.$$

Both these equations reduce to the one equation $2x - y = 0$. But we know that the eigenvector $\begin{pmatrix} x \\ y \end{pmatrix}$ is a probability vector so that $x + y = 1$.

Hence from $2x - y = 0$ and $x + y = 1$, $x = \frac{1}{3}$, $y = \frac{2}{3}$. For the other eigenvector we know $\lambda = -\frac{1}{5}$ (check $a + b - 1 = \frac{1}{5} + \frac{3}{5} - 1 = -\frac{1}{5}$) then

$$\begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -\frac{1}{5} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Hence

$$\frac{1}{5}x + \frac{2}{5}y = -\frac{1}{5}x$$

and

$$\frac{4}{5}x + \frac{3}{5}y = -\frac{1}{5}y.$$

$$\frac{2}{5}x + \frac{2}{5}y = 0$$

and

$$\frac{4}{5}x + \frac{4}{5}y = 0.$$

Both these equations reduce to $x + y = 0$. But since $\begin{pmatrix} x \\ y \end{pmatrix}$ is a probability vector, $x + y = 1$ - so there is no solution for x and y between these equations and the only suitable vector is $\begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \end{pmatrix}$. The stable state is given when $\lambda = 1$ and then

$$\begin{pmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

(check with earlier result).

Example

8.11 Show from the general 2×2 matrix $\begin{pmatrix} a & 1-b \\ 1-a & b \end{pmatrix}$ that the stable-state eigenvector ($\lambda = 1$) is

$$\begin{pmatrix} \frac{b-1}{a+b-2} \\ \frac{a-1}{a+b-2} \end{pmatrix}.$$

Consider the examples in the text by this easier method of finding the stable-state matrix which the situation will tend towards.

8.7 Three outcome situations

Situations giving 3×3 transition matrices may also be investigated in this way. In example (ii) (page 98) concerning choice of the nut, raisin and plain bars of chocolate, what is the probability of having a nut bar on the last day of term?

The transition matrix is

$$\begin{array}{ccccc} & N & R & P & \text{1st choice} \\ \text{2nd choice} & N & R & P & \\ & \begin{pmatrix} 0 & \frac{1}{3} & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{2}{3} & 0 \end{pmatrix} & & & \end{array}$$

and are seeking the 'steady-state' eigenvector (again assuming the

term is endless). If $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is this eigenvector then $\begin{pmatrix} 0 & \frac{1}{3} & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

and also $x + y + z = 1$.

We have the equations

$$\frac{1}{3}y + z = x \quad . \quad . \quad . \quad . \quad (1)$$

$$\frac{1}{2}x = y \quad . \quad . \quad . \quad . \quad (2)$$

$$\frac{1}{2}x + \frac{2}{3}y = z \quad . \quad . \quad . \quad . \quad (3)$$

$$x + y + z = 1 \quad . \quad . \quad . \quad . \quad (4).$$

From (2) $x = 2y$ and then in (1) $3z = 5y$ (which is the same equation as is obtained by putting $x = 2y$ in (3)).

Then from (4)

$$2y + y + 5y/3 = 1$$

and

$$\text{Hence} \quad x = \frac{6}{14} \quad y = \frac{3}{14} \quad \text{and} \quad z = \frac{5}{14}.$$

At the end of term the 'steady-state' matrix is

$$\begin{pmatrix} \frac{6}{14} & \frac{6}{14} & \frac{6}{14} \\ \frac{3}{14} & \frac{3}{14} & \frac{3}{14} \\ \frac{5}{14} & \frac{5}{14} & \frac{5}{14} \end{pmatrix}$$

and as a nut bar was chosen at the beginning of term, the end of term probabilities are given by

$$\begin{pmatrix} \frac{6}{14} & \frac{6}{14} & \frac{6}{14} \\ \frac{3}{14} & \frac{3}{14} & \frac{3}{14} \\ \frac{5}{14} & \frac{5}{14} & \frac{5}{14} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{6}{14} \\ \frac{3}{14} \\ \frac{5}{14} \end{pmatrix} \begin{matrix} N \\ R \\ P \end{matrix}$$

that is a probability of $\frac{3}{7}$ of ending the term with a nut bar. These also show the proportion of nut, raisin and plain bars of chocolate I have had during the term.

Examples

8.12 A lecturer is either late, early or on-time for his lectures. If he is late for a lecture, he is always early for the next. However, if he is

early it is twice as likely he will be late for the next lecture as he will be on-time. If he is on-time then it is twice as likely he will be late as early for the next lecture. If he is early for his first lecture of term what are the probabilities of his arrival states for the end of term lecture?

- 8.13** The diagram shows four compartments with doors leading from one to another. A mouse in any compartment is equally likely to pass through each of the doors of the compartment in going to the next. Find the transition matrix and the probabilities of the mouse being in the various compartments in the long run.

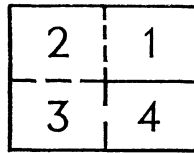


Fig. 8.1

8.8 Absorbing state situations

The examples looked at have been of chain situations in which the successive application of the transition matrix has shown the outcome of the chain at any stage. Such situations are examples of *ergodic chains* or *Markov chains*. Since the chains continue for as long as we please they are chains with *non-absorbing states*. An absorbing state is a situation in which nothing further can happen and the chain ends if that state is obtained. Examples with absorbing states are more difficult to deal with but are often more interesting. A situation of a Markov chain with absorbing states is the following.

A 'drunk' lives on a house-boat and has to negotiate the gang-plank each evening. One pace to the left or right gets him in the water (the absorbing state!!). He wants to get home but the probability (after any pace) that he takes a pace forward is $\frac{1}{3}$, while there are probabilities of $\frac{1}{6}$ that he will take a pace backwards, $\frac{1}{4}$ he will stagger left into the water and $\frac{1}{4}$ he staggers right. (An example of this kind is known as a Random Walk!)

The transition matrix for one step to the next is given by

$$P = \begin{matrix} & \begin{matrix} F & B & L & R \end{matrix} \text{ from} \\ \begin{matrix} F \\ B \\ L \\ R \end{matrix} \text{ to} & \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 1 \end{pmatrix} \end{matrix}$$

The positions L and R are absorbing states, since once in them it is impossible to leave them whilst having paced forwards or backwards it is possible to move into any of the other states.

The drunk is just on the gang-plank so the situation is

$$\begin{matrix} F \\ B \\ L \\ R \end{matrix} \begin{pmatrix} \frac{1}{3} \\ \frac{1}{6} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}$$

for the first step. For the second step:

$$\begin{aligned} L\left(\frac{1}{4}\right) &\longrightarrow 1 \longrightarrow L\left(\frac{1}{4}\right) \\ R\left(\frac{1}{4}\right) &\longrightarrow 1 \longrightarrow R\left(\frac{1}{4}\right) \\ F\left(\frac{1}{3}\right) &\begin{cases} \xrightarrow{\frac{1}{4}} L\left(\frac{1}{12}\right) \\ \xrightarrow{\frac{1}{4}} R\left(\frac{1}{12}\right) \\ \xrightarrow{\frac{1}{3}} F\left(\frac{1}{9}\right) \\ \xrightarrow{\frac{1}{6}} B\left(\frac{1}{18}\right) \end{cases} \\ B\left(\frac{1}{6}\right) &\begin{cases} \xrightarrow{\frac{1}{4}} L\left(\frac{1}{24}\right) \\ \xrightarrow{\frac{1}{4}} R\left(\frac{1}{24}\right) \\ \xrightarrow{\frac{1}{3}} F\left(\frac{1}{18}\right) \\ \xrightarrow{\frac{1}{6}} B\left(\frac{1}{36}\right) \end{cases} \end{aligned}$$

So the probability of 2 steps forward is $\frac{1}{9} + \frac{1}{18} = \frac{1}{6}$. Note that once in L or R that state continues. The probabilities for the second step come

from using the starting probability vector $\begin{pmatrix} \frac{1}{3} \\ \frac{1}{6} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}$ with the transition

matrix as

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{1}{6} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ \frac{1}{12} \\ \frac{3}{8} \\ \frac{3}{8} \end{pmatrix} \begin{matrix} F \\ B \\ L \\ R \end{matrix}.$$

Again the matrix easily picks up the probabilities. Note that if the drunk went into the water on the first step then the chain finishes – but we assume it to continue – letting the drunk bob up and down in the water for each event considered! Again we need to take powers of the transition matrix to be able to describe the chain as it continues. The form of $\mathbf{P}$ allows this arithmetic to be relatively easy. Since $\mathbf{P}$ has absorbing states (the 1 and 0 columns) it can always be partitioned into *canonical* form as

$$\begin{matrix} r & s \\ r & s \\ s \end{matrix} \left(\begin{array}{c|c} \mathbf{Q} & \mathbf{O} \\ \hline \mathbf{R} & \mathbf{I} \end{array} \right).$$

Where $\mathbf{I}$ is an $s \times s$ identity matrix, $\mathbf{O}$ is an $r \times r$ zero matrix, $\mathbf{R}$ and $s \times r$ matrix and $\mathbf{Q}$ an $r \times r$ matrix. Is it possible to treat the matrix as

$$\begin{pmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{R} & \mathbf{I} \end{pmatrix} \text{ and multiply it by itself just as we do 'true' matrices? Does } \begin{pmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{R} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{R} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{Q} \cdot \mathbf{Q} + \mathbf{O} \cdot \mathbf{R} & \mathbf{Q} \cdot \mathbf{O} + \mathbf{O} \cdot \mathbf{I} \\ \mathbf{R} \cdot \mathbf{Q} + \mathbf{I} \cdot \mathbf{R} & \mathbf{R} \cdot \mathbf{O} + \mathbf{I}^2 \end{pmatrix} = \begin{pmatrix} \mathbf{Q}^2 & \mathbf{O} \\ \mathbf{R} + \mathbf{RQ} & \mathbf{I}^2 \end{pmatrix}?$$

The reader is invited to verify that the partitioning of matrices and multiplying them matrix-wise as blocks does work. (Use two 3×3 matrices as

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} j & k & l \\ m & n & o \\ p & q & r \end{pmatrix}$$

and obtain the result.)

Then partition them as

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} j & k & l \\ m & n & o \\ p & q & r \end{pmatrix}$$

and let

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix} = \mathbf{A}; \quad \begin{pmatrix} c \\ f \end{pmatrix} = \mathbf{B}; \quad (g \ h) = \mathbf{C}; \quad (i) = \mathbf{D}; \quad \begin{pmatrix} j & k \\ m & n \end{pmatrix} = \mathbf{E}; \\ \begin{pmatrix} l \\ o \end{pmatrix} = \mathbf{F}; \quad (p \ q) = \mathbf{G} \quad \text{and} \quad (r) = \mathbf{H}.$$

Multiply $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix}$ as matrices and verify, when the capitals are reconstituted as matrices and multiplied out (care with the right order of the capitals) that the results are the same.

As

$$\begin{pmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{R} & \mathbf{I} \end{pmatrix}^2 = \begin{pmatrix} \mathbf{Q}^2 & \mathbf{O} \\ \mathbf{R} + \mathbf{RQ} & \mathbf{I} \end{pmatrix}, \\ \begin{pmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{R} & \mathbf{I} \end{pmatrix}^3 = \begin{pmatrix} \mathbf{Q}^2 & \mathbf{O} \\ \mathbf{R} + \mathbf{RQ} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{R} & \mathbf{I} \end{pmatrix} \\ = \begin{pmatrix} \mathbf{Q}^3 & \mathbf{O} \\ \mathbf{R} + \mathbf{RQ} + \mathbf{RQ}^2 & \mathbf{I} \end{pmatrix}$$

and by induction

$$\begin{pmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{R} & \mathbf{I} \end{pmatrix}^n = \begin{pmatrix} \mathbf{Q}^n & \mathbf{O} \\ \mathbf{R} + \mathbf{RQ} + \mathbf{RQ}^2 + \dots + \mathbf{RQ}^{n-1} & \mathbf{I} \end{pmatrix}.$$

Also

$$\begin{aligned} \mathbf{R} + \mathbf{RQ} + \mathbf{RQ}^2 + \dots + \mathbf{RQ}^{n-1} &= \mathbf{R}(\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \dots + \mathbf{Q}^{n-1}) \\ &= \mathbf{R} \left(\frac{\mathbf{I} - \mathbf{Q}^n}{\mathbf{I} - \mathbf{Q}} \right). \end{aligned}$$

In our example $\mathbf{R} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix}$ and $\mathbf{Q} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} \end{pmatrix}$.

If we allow the drunk to be able to take an infinity of steps $\mathbf{Q}^n \rightarrow 0$ as $n \rightarrow \infty$. (Check this by finding $\mathbf{Q}^n$ – the eigenvalues of $\mathbf{Q}$ are 0 and $\frac{1}{2}$ and hence $\mathbf{Q}^n = \mathbf{P}^{-1}\mathbf{D}^n\mathbf{P}$ and $\mathbf{D}^n$ will be $\begin{pmatrix} 0 & 0 \\ 0 & (\frac{1}{2})^n \end{pmatrix}$ which $\rightarrow 0$ as $n \rightarrow \infty$.)

The matrix $\mathbf{Q}$ describes the non-absorbing states so that entries in $\mathbf{Q}^n$ describe the probability of being in any one of these non-absorbing states after n steps. Hence $\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \mathbf{Q}^3 + \dots$ gives the sum of the probabilities of being in a non-absorbing state as n increases.

As $\mathbf{Q}^n \rightarrow 0$ as $n \rightarrow \infty$,

$$\begin{pmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{R} & \mathbf{I} \end{pmatrix}^n \rightarrow \begin{pmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{R}(\mathbf{I} - \mathbf{Q})^{-1} & \mathbf{I} \end{pmatrix} \quad \text{as } n \rightarrow \infty$$

since

$$\mathbf{R} \left(\frac{\mathbf{I} - \mathbf{Q}^n}{\mathbf{I} - \mathbf{Q}} \right) \rightarrow \mathbf{R}(\mathbf{I} - \mathbf{Q})^{-1} \quad \text{as } n \rightarrow \infty.$$

Now

$$\mathbf{I} - \mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{5}{6} \end{pmatrix} \quad \text{and} \quad (\mathbf{I} - \mathbf{Q})^{-1} = 2 \begin{pmatrix} \frac{5}{6} & \frac{1}{3} \\ \frac{1}{6} & \frac{2}{3} \end{pmatrix}$$

and hence

$$\mathbf{R}(\mathbf{I} - \mathbf{Q})^{-1} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix} 2 \begin{pmatrix} \frac{5}{6} & \frac{1}{3} \\ \frac{1}{6} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

After an infinity of paces the transition matrix becomes

$$\begin{array}{ccccc} & F & B & L & R \\ \begin{array}{l} F \\ B \\ L \\ R \end{array} & \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{pmatrix} \end{array}$$

showing that the probability of absorption in each of the absorbing states L and R is each $\frac{1}{2}$ starting from either of the non-absorbing states.

8.9 A Game example

Many games involve probability and absorbing states – winning or losing – and such situations can be considered by matrix methods.

A simplified game of snakes and ladders could use such a board as

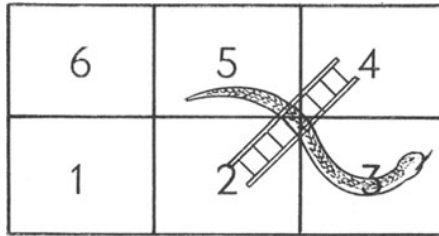


Fig. 8.2

with rules of up ladder and down snake; no throw of a 6 on the dice to start; if too large a number is thrown to land on 6, 'home', the player remains on that square and once 'home' on 6 the player remains there (the absorbing state).

We can find the transition matrix of probabilities of moving from one square to another square, as

$$M = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} & \text{from} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \text{to} & \begin{pmatrix} \frac{1}{6} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{3} & \frac{1}{6} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{6} & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & 0 & 1 \end{pmatrix} \end{matrix}$$

where entries have been worked out, for example: being on square 3 and remaining on 3 (column 3, row 3), that is a throw with the dice keeps the counter on 3; as a throw of 1, counter goes to 4; a throw of 2, counter on 3 (down snake); throw of 3, counter on 6; throws of 4, 5, 6 are too high and counter remains on 3. So 4 throws from 6 keeps the counter on 3 and then the probability of being on 3 and remaining on 3 is $\frac{4}{6} = \frac{2}{3}$. Also we see the probability of being on 3 and going to 4 is $\frac{1}{6}$; to 5, 0 and to 6, $\frac{1}{6}$. One does not go back to squares 2 or 1.

The probabilities of being on any one of the 6 squares after the first throw are given by

$$M \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = M \mathbf{x} = \begin{pmatrix} \frac{1}{6} \\ 0 \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \\ \frac{1}{6} \end{pmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix}$$

so the probability of being on square 6 and home after one throw is $\frac{1}{6}$.

The probabilities after the second throw of the dice are given by

$$\mathbf{M}^2 = \text{to } \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} (\frac{1}{6})^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{36} & \frac{2}{9} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{2}{9} & \frac{1}{36} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{11}{36} & 0 & \frac{11}{36} & \frac{11}{36} & 0 & 1 \end{pmatrix} \end{matrix} \text{ from}$$

The probabilities in $\mathbf{M}^2$ can be read off, for example, as the probability of being on 3 (column 3) and remaining on 3 (row 3) after two throws of the dice is $\frac{1}{36}$; the probability of being on 3 and moving to 4 after two throws is $\frac{2}{9}$. It is a useful exercise at this stage to verify these probabilities by the 'tree' method used earlier – then the usefulness of the matrix method becomes apparent; it is a direct and explicit routine method.

The probability of being on any square after two throws is given by

$$\mathbf{M}^2 \mathbf{x} = \begin{pmatrix} \frac{1}{36} \\ 0 \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \\ \frac{11}{36} \end{pmatrix}; \text{ and the probability of being home after two throws is } \frac{11}{36}.$$

This probability includes that of being home after 1 throw plus probability of being home after the next throw. To find the probability of being home in precisely two throws we subtract the probability of being home in 1 throw ($\frac{1}{6}$) from $\frac{11}{36}$. The probability of being home in 2 throws exactly is given by the entry at the bottom of the column vector $(\mathbf{M}^2 - \mathbf{M})\mathbf{x}$. Continuing; the probability of being home in exactly 3 throws is the entry at the bottom of the column vector $(\mathbf{M}^3 - \mathbf{M}^2)\mathbf{x}$ etc.

The average length of a game is considered as

$$1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 + \dots$$

where p_n is the probability of being home in exactly n throws.

For this situation the average length of the game is given by

$$1 \cdot \mathbf{M}\mathbf{x} + 2 \cdot (\mathbf{M}^2 - \mathbf{M})\mathbf{x} + 3(\mathbf{M}^3 - \mathbf{M}^2)\mathbf{x} + \dots + n(\mathbf{M}^n - \mathbf{M}^{n-1})\mathbf{x} \quad (1)$$

when $n \rightarrow \infty$, considering the entry at the bottom of the column vector.

Now $\mathbf{M}$ can be partitioned as

$$\begin{matrix} & \begin{matrix} 1 & 3 & 4 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 3 \\ 4 \\ 6 \end{matrix} & \begin{pmatrix} \frac{1}{6} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{6} & 0 \\ \frac{1}{3} & \frac{1}{6} & \frac{2}{3} & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 1 \end{pmatrix} \end{matrix}$$

and we have $\mathbf{Q}$ as $\begin{pmatrix} \frac{1}{6} & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{2}{3} \end{pmatrix}$ and $\mathbf{R}$ as $\begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$, so $\mathbf{M} = \begin{pmatrix} \mathbf{Q} & \mathbf{O} \\ \mathbf{R} & \mathbf{I} \end{pmatrix}$. As

before as we take powers of $\mathbf{M}$, from page (108),

$$\mathbf{M}^n = \begin{pmatrix} \mathbf{Q}^n & \mathbf{O} \\ \mathbf{R} \frac{(\mathbf{I} - \mathbf{Q}^n)}{(\mathbf{I} - \mathbf{Q})} & \mathbf{I} \end{pmatrix}.$$

Thus

$$\begin{aligned} \mathbf{M}^n - \mathbf{M}^{n-1} &= \begin{pmatrix} \mathbf{Q}^n & \mathbf{O} \\ \mathbf{R} \frac{(\mathbf{I} - \mathbf{Q}^n)}{(\mathbf{I} - \mathbf{Q})} & \mathbf{I} \end{pmatrix} - \begin{pmatrix} \mathbf{Q}^{n-1} & \mathbf{O} \\ \mathbf{R} \frac{(\mathbf{I} - \mathbf{Q}^{n-1})}{(\mathbf{I} - \mathbf{Q})} & \mathbf{I} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{Q}^n - \mathbf{Q}^{n-1} & \mathbf{O} \\ -\frac{\mathbf{R}}{\mathbf{I} - \mathbf{Q}} (\mathbf{Q}^n - \mathbf{Q}^{n-1}) & \mathbf{O} \end{pmatrix} \\ &= (\mathbf{Q}^n - \mathbf{Q}^{n-1}) \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\frac{\mathbf{R}}{\mathbf{I} - \mathbf{Q}} & \mathbf{O} \end{pmatrix} \end{aligned}$$

Substituting for values of n from 1 to n in expression (1) leaving the $\mathbf{x}$ as a factor, the average length of a game is given by

$$\begin{aligned} &\left(1 \cdot (\mathbf{Q} - \mathbf{I}) + 2(\mathbf{Q}^2 - \mathbf{Q}) + 3(\mathbf{Q}^3 - \mathbf{Q}^2) + \dots + n(\mathbf{Q}^n - \mathbf{Q}^{n-1}) \right) \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\frac{\mathbf{R}}{\mathbf{I} - \mathbf{Q}} & \mathbf{O} \end{pmatrix} \mathbf{x} \\ &= \left(-\mathbf{I} - \mathbf{Q} - \mathbf{Q}^2 - \mathbf{Q}^3 \dots - \mathbf{Q}^{n-1} + n\mathbf{Q}^{n-1} \right) \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\frac{\mathbf{R}}{\mathbf{I} - \mathbf{Q}} & \mathbf{O} \end{pmatrix} \mathbf{x} \\ &= \left(-\frac{(\mathbf{I} - \mathbf{Q}^n)}{\mathbf{I} - \mathbf{Q}} + n\mathbf{Q}^{n-1} \right) \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\frac{\mathbf{R}}{\mathbf{I} - \mathbf{Q}} & \mathbf{O} \end{pmatrix} \mathbf{x}. \end{aligned}$$

It can be shown for the matrix $\mathbf{Q}$ that $\mathbf{Q}^n \rightarrow 0$ as $n \rightarrow \infty$ and also $n\mathbf{Q}^{n-1} \rightarrow 0$, so the average length of game becomes

$$\begin{pmatrix} -\mathbf{I} \\ \mathbf{I} - \mathbf{Q} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{O} \\ -\frac{\mathbf{R}}{\mathbf{I} - \mathbf{Q}} & \mathbf{O} \end{pmatrix} \mathbf{x} = \begin{pmatrix} -\frac{\mathbf{I}}{\mathbf{I} - \mathbf{Q}} & \mathbf{O} \\ \frac{\mathbf{R}}{(\mathbf{I} - \mathbf{Q})^2} & \mathbf{O} \end{pmatrix} \mathbf{x}.$$

$$Q = \begin{pmatrix} \frac{1}{6} & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{2}{3} \end{pmatrix}, \text{ so } I - Q = \begin{pmatrix} \frac{5}{6} & 0 & 0 \\ -\frac{1}{3} & \frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{3} & -\frac{1}{6} & \frac{1}{3} \end{pmatrix}$$

and $(I - Q)^{-1}$ becomes $\begin{pmatrix} \frac{6}{5} & 0 & 0 \\ \frac{12}{5} & 4 & 2 \\ \frac{12}{5} & 2 & 4 \end{pmatrix}.$

$$\begin{aligned} \frac{R}{(I - Q)^2} &= \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} \frac{6}{5} & 0 & 0 \\ \frac{12}{5} & 4 & 2 \\ \frac{12}{5} & 2 & 4 \end{pmatrix} \begin{pmatrix} \frac{6}{5} & 0 & 0 \\ \frac{12}{5} & 4 & 2 \\ \frac{12}{5} & 2 & 4 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} \frac{36}{25} & 0 & 0 \\ \frac{432}{25} & 20 & 20 \\ \frac{432}{25} & 20 & 20 \end{pmatrix} = \begin{pmatrix} 6 & \frac{40}{6} & \frac{40}{6} \end{pmatrix}. \end{aligned}$$

We thus have $\frac{-I}{I - Q}$ as a 3×3 matrix and the average length of a game as

$$\left(\begin{array}{ccc|c} -I & & & \\ \hline I - Q & & & O \\ \hline 6 & \frac{40}{6} & \frac{40}{6} & O \end{array} \right) \mathbf{x}$$

and the entry required is $6 -$ the average length of the game.

With the previous example of the drunkard's walk, we can find the average number of paces before he falls into the absorbing states. Here

$R \left(\frac{I}{(I - Q)^2} \right)$ gives $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ so the drunkard, on the average, takes only two paces before getting wet.

As an exercise on the work covered here; find the average length of path of the snakes and ladders game when the snake is removed; when the ladder is removed; when the snake and ladder are removed. The answer in each case is 6, since the 6-sided dice can always finish the game from any position on the 6-square board. But consider the situation on a 9-square board.

The situations given here are intended to indicate the scope of matrix methods. There are other means of dealing with these situations and obtaining the answers found, but the application of matrices describes the situation neatly and readily produces the answers. The technique of matrix multiplication turns up surprisingly often.

QUADRATIC FORMS

9.1 Quadratic forms in terms of matrices

If $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ then $\mathbf{x}' = (x \ y)$ and $\mathbf{x}'\mathbf{x} = (x \ y) \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2$, that is the circle $x^2 + y^2 = a^2$ can be represented in matrix form by

$$(x \ y) \begin{pmatrix} x \\ y \end{pmatrix} = a^2$$

or

$$(x \ y) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = a^2.$$

Similarly the equation $\mathbf{x}' \begin{pmatrix} 6 & 1 \\ 3 & 3 \end{pmatrix} \mathbf{x} = 14$ reduces as follows

$$(x \ y) \begin{pmatrix} 6 & 1 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 14 \quad . \quad . \quad . \quad (1),$$

$$(x \ y) \begin{pmatrix} 6x+y \\ 3x+3y \end{pmatrix} = 14$$

$$6x^2 + xy + 3xy + 3y^2 = 14$$

and $6x^2 + 4xy + 3y^2 = 14 \quad . \quad . \quad . \quad (2).$

The matrix notation used in (1) for the quadratic equation (2) shows another use of matrices. Scalars, vectors and matrices can be regarded as successively more complicated types of number, each having a variety of applications. Just as ordinary (scalar) numbers can be used to describe the size of a class or the temperature of a room, and a vector to describe the position of a point, or the direction of a line, so a matrix may be used to describe a rotation of axes or a quadratic form. The use of the matrix to describe a quadratic form has wide application in geometry and its importance lies in its easy extension into 3 or n dimensions.

Examples

9.1 What is the equation represented by $\mathbf{x}'\mathbf{A}\mathbf{x} = 1$ when

$$\mathbf{A} = \begin{pmatrix} 6 & 1 \\ 3 & 3 \end{pmatrix}; \quad \begin{pmatrix} 6 & -2 \\ 6 & 3 \end{pmatrix}; \quad \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix}?$$

9.2 Can you find another matrix $\mathbf{A}$ to give the equation (2) above?

9.3 Put the equations with matrix notation

$$(i) \quad 3x^2 + 2xy + y^2 = 1$$

$$(ii) \quad 4x^2 - 3xy + 2y^2 = 1.$$

9.2 Reduction to Diagonal Form

The general homogeneous quadratic form $ax^2 + 2hxy + by^2 = c$ can be represented in matrix form as $\mathbf{x}' \begin{pmatrix} a & h \\ h & b \end{pmatrix} \mathbf{x} = c$ where the matrix $\begin{pmatrix} a & h \\ h & b \end{pmatrix}$ has been chosen here as a *symmetric* matrix – the most useful form, as will be seen from later work.

This quadratic form represents a conic section, which will be a pair of straight lines, a circle, an ellipse or a hyperbola depending on the values of a , b and h . The centre of the conic is the origin but its axes do not necessarily lie along the coordinate axes. It is possible to find the direction of the axes of the conic and its form, by finding the eigenvector of the corresponding matrix.

Consider the quadratic form $6x^2 + 4xy + 3y^2 = 14$. When written in matrix form $\mathbf{x}'\mathbf{A}\mathbf{x} = 14$, with $\mathbf{A}$ a symmetric matrix, we have $\mathbf{A} = \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix}$. The eigenvalues of $\mathbf{A}$ are given by

$$(6 - \lambda)(3 - \lambda) - 4 = 0$$

$$\text{i.e.} \quad \lambda^2 - 9\lambda + 14 = 0$$

$$\text{i.e.} \quad (\lambda - 7)(\lambda - 2) = 0$$

$$\text{and} \quad \lambda = 2 \text{ or } 7.$$

When $\lambda = 2$, from $(6 - \lambda)x + 2y = 0$ the eigenvector is given by $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

When $\lambda = 7$, the eigenvector is given by $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ (perpendicular to the first eigenvector, as $\mathbf{A}$ is symmetric).

Considering the eigenvector with positive gradient, $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$; if we make this an axis, then the matrix which will rotate the axes to this position is given by $\mathbf{R} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$ where $\tan \alpha = \frac{1}{2}$, as for the eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$; $x = 2$, $y = 1$; so $\mathbf{R} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$. What will happen to the quadratic form when we refer it to axes taken with respect to the eigenvector

of the matrix? If the new axes are labelled x_1, y_1 then the rotation matrix has the effect $\mathbf{x}_1 = \mathbf{R}\mathbf{x}$ where (x_1, y_1) are the coordinates of the point (x, y) when referred to the new axes.

From $\mathbf{x}_1 = \mathbf{R}\mathbf{x}$ then $\mathbf{x} = \mathbf{R}^{-1}\mathbf{x}_1$.

Hence the original quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x} = 14$ becomes

$$(\mathbf{R}^{-1}\mathbf{x}_1)'\mathbf{A}(\mathbf{R}^{-1}\mathbf{x}_1) = 14.$$

Hence $\mathbf{x}_1'(\mathbf{R}^{-1})'\mathbf{A}\mathbf{R}^{-1}\mathbf{x}_1 = 14$ as $(\mathbf{R}^{-1}\mathbf{x}_1)' = \mathbf{x}_1'(\mathbf{R}^{-1})'$ or $\mathbf{x}_1'\mathbf{R}\mathbf{A}\mathbf{R}'\mathbf{x}_1 = 14$ as $\mathbf{R}^{-1} = \mathbf{R}'$ and $(\mathbf{R}')' = \mathbf{R}$. The quadratic form considered thus becomes

$$\mathbf{x}_1' \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \mathbf{x}_1 = 14.$$

The matrices reduce to $\begin{pmatrix} 7 & 0 \\ 0 & 2 \end{pmatrix}$ – a matrix which is said to be of *diagonal* form, that is with entries only in the leading diagonal, other values being zero. The values of the entries in the leading diagonal of this matrix are the eigenvalues of the original matrix.

The quadratic form has been reduced to

$$\mathbf{x}_1' \begin{pmatrix} 7 & 0 \\ 0 & 2 \end{pmatrix} \mathbf{x}_1 = 14$$

or

$$(x_1 \ y_1) \begin{pmatrix} 7 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = 14,$$

$$7x_1^2 + 2y_1^2 = 14,$$

or

$$\frac{x_1^2}{2} + \frac{y_1^2}{1} = 1$$

an ellipse with semi major axis of $\sqrt{2}$ and semi minor axis of $\sqrt{1}$.

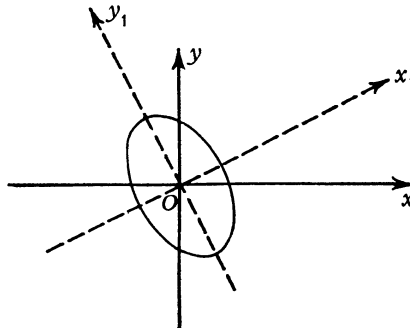


Fig. 9.1

9.3 Worked investigations

(i) Interpret the form of the quadratic equation $4x^2 - 12xy + 9y^2 = 13$.

The quadratic form in matrix notation is $\mathbf{x}' \begin{pmatrix} 4 & -6 \\ -6 & 9 \end{pmatrix} \mathbf{x} = 13$ where the matrix is symmetric. The eigenvalues of $\begin{pmatrix} 4 & -6 \\ -6 & 9 \end{pmatrix}$ are given by $(4 - \lambda)(9 - \lambda) - 36 = 0$

$$\lambda^2 - 13\lambda = 0$$

$$\lambda = 0 \text{ or } 13.$$

The eigenvector when $\lambda = 0$, from $(4 - \lambda)x - 6y = 0$ is $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ and the eigenvector when $\lambda = 13$ is $\begin{pmatrix} 2 \\ -3 \end{pmatrix}$.

Considering axes rotated so that the eigenvector $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ becomes the new x_1 axis; the rotation matrix is given by

$$\frac{1}{\sqrt{13}} \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix}, \text{ as } \tan \alpha = \frac{2}{3}$$

and thus the quadratic form when referred to $x_1 y_1$ axes becomes

$$\mathbf{x}'_1 \frac{1}{\sqrt{13}} \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 4 & -6 \\ -6 & 9 \end{pmatrix} \frac{1}{\sqrt{13}} \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \mathbf{x}_1 = 13.$$

The matrices reduce to $\begin{pmatrix} 0 & 0 \\ 0 & 13 \end{pmatrix}$ - again of diagonal form and with diagonal entries equal to the eigenvalues.

The quadratic form now is

$$\mathbf{x}' \begin{pmatrix} 0 & 0 \\ 0 & 13 \end{pmatrix} \mathbf{x} = 13$$

or

$$(x_1 \ y_1) \begin{pmatrix} 0 & 0 \\ 0 & 13 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = 13,$$

$$13y_1^2 = 13,$$

or

$$y_1^2 = 1$$

and

$$y_1 = \pm 1$$

which represents two straight lines.

(The original equation was $4x^2 - 12xy + 9y^2 = 13$ giving $2x - 3y = +\sqrt{13}$ or $2x - 3y = -\sqrt{13}$; two straight lines.)

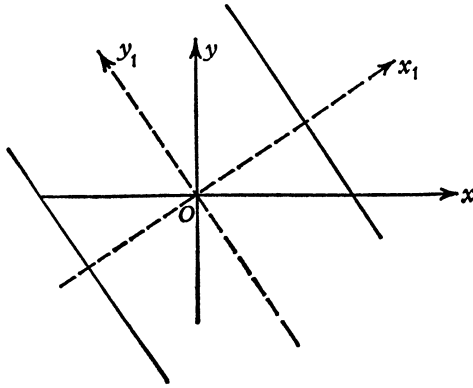


Fig. 9.2

Note also that the vanishing of the determinant of the matrix, which implies an eigenvector is zero, is also the condition for the second degree terms to form a perfect square.

(ii) Interpret the form of the quadratic equation

$$x^2 - 8xy + 7y^2 = 9.$$

In matrix notation this becomes

$$\mathbf{x}' \begin{pmatrix} 1 & -4 \\ -4 & 7 \end{pmatrix} \mathbf{x} = 9.$$

The eigenvalues are given by

$$(1 - \lambda)(7 - \lambda) - 16 = 0$$

$$7 - 8\lambda + \lambda^2 - 16 = 0$$

$$\lambda^2 - 8\lambda - 9 = 0$$

$$(\lambda - 9)(\lambda + 1) = 0,$$

therefore

$$\lambda = 9 \text{ or } -1.$$

The eigenvector, when $\lambda = 9$, from $(1 - \lambda)x - 4y = 0$ is $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ and when $\lambda = -1$ the eigenvector is $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Taking the eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, of gradient $\frac{1}{2}$ the rotation matrix is $\frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$.

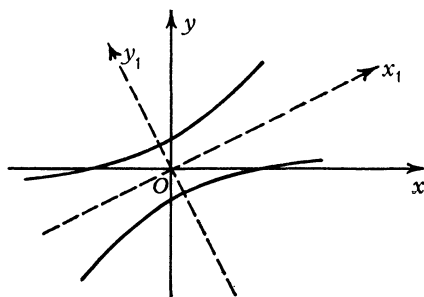


Fig. 9.3

The quadratic form when referred to the eigenvector as new x_1 axis is given by

$$\mathbf{x}'_1 \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -4 \\ -4 & 7 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \mathbf{x}_1 = 9$$

or

$$\mathbf{x}_2 \begin{pmatrix} -1 & 0 \\ 0 & 9 \end{pmatrix} \mathbf{x}_2 = 9,$$

$$-x_1^2 + 9y_1^2 = 9,$$

$$-\frac{x_1^2}{9} + y_1^2 = 1, \quad \text{a hyperbola.}$$

Examples

Investigate and sketch the following quadratic forms as conics.

9.4 $3x^2 - 2xy + 3y^2 = 8.$

9.5 $x^2 + 4xy - 2y^2 = 6.$

9.6 $4x^2 - 4xy + y^2 = 20.$

9.7 $4x^2 + 4y^2 = 9.$

An equation of the form $ax^2 + 2hxy + by^2 = c$ cannot be a parabola as a change in sign of x and y leaves the equation unaltered, and hence it represents a curve symmetrical about the origin.

9.4 Formulation and proof of general results

In the examples which have been worked through, several key points always appear. These need to be looked at generally and proofs of them established.

The quadratic form used viz: $6x^2 + 4xy + 3y^2 = 14$ was put into matrix notation as $\mathbf{x}' \begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix} \mathbf{x} = 14$ with the *symmetric* matrix $\begin{pmatrix} 6 & 2 \\ 2 & 3 \end{pmatrix}$. The eigenvalues of this matrix were found to be 2 and 7 and the corresponding eigenvectors $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. When the quadratic form was referred to the eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ as axis, its resultant form became $\begin{pmatrix} 7 & 0 \\ 0 & 2 \end{pmatrix} - a$ *matrix of diagonal form with entries in the leading diagonal of the eigenvalues*, with the eigenvalue 7 (corresponding to the eigenvector used) in the top left hand corner. Will these results be true generally?

Consider a symmetric matrix $\mathbf{A}$ which has eigenvalues of λ_1 and λ_2 and which will have eigenvectors which are orthogonal. If we choose one of the eigenvectors to be an axis and rotate the axes according to the rotation of axes matrix $\mathbf{R} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$, then the matrix $\mathbf{A}$ transforms into $\mathbf{RAR}'$ (see page 57).

The eigenvalues of $\mathbf{RAR}'$ have been found to be λ_1 and λ_2 and $\mathbf{RAR}'$ is of the diagonal form $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

The eigenvalues of $\mathbf{RAR}'$ are given by the roots of the equation $\det(\mathbf{RAR}' - \lambda\mathbf{I}) = 0$.

But $\mathbf{RAR}' - \lambda\mathbf{I} = \mathbf{RAR}' - \mathbf{RAIR}'$ since $\mathbf{RI} = \mathbf{IR}$ and $\mathbf{RR}' = \mathbf{RR}^{-1} = \mathbf{I}$

$= \mathbf{R}(\mathbf{AR}' - \lambda\mathbf{IR}') \text{ by the distributive law}$

$= \mathbf{R}(\mathbf{A} - \lambda\mathbf{I})\mathbf{R}' \text{ by the distributive law}$

therefore

$$\det(\mathbf{RAR}' - \lambda\mathbf{I}) = \det \mathbf{R} \det(\mathbf{A} - \lambda\mathbf{I}) \det \mathbf{R}' \quad (\text{page 30})$$

$$= \det(\mathbf{A} - \lambda\mathbf{I}) \text{ since } \det \mathbf{R} = \det \mathbf{R}' = 1$$

that is $\mathbf{A}$ and $\mathbf{RAR}'$ have the same characteristic equations. These equations must therefore have the same roots, and so have the same eigenvalues.

To prove $\mathbf{RAR}' = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, let one eigenvector be given by the unit vector $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ and the other eigenvector by the unit vector $\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$. Then $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ may be written as $\begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$ for some α and $\begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ will be represented by $\begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}$ since $\mathbf{c}$ and $\mathbf{d}$ are orthogonal as $\mathbf{A}$ is symmetric (page 53).

Then $\mathbf{c}'\mathbf{c} = (c_1 c_2) \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = c_1^2 + c_2^2 = \cos^2 \alpha + \sin^2 \alpha = 1$ and similarly

$$\mathbf{d}'\mathbf{d} = d_1^2 + d_2^2 = 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$\mathbf{c}'\mathbf{d} = (c_1 c_2) \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = c_1 d_1 + c_2 d_2 = \cos \alpha \cdot -\sin \alpha + \sin \alpha \cdot \cos \alpha = 0 = \mathbf{d}'\mathbf{c} \quad (2)$$

$\mathbf{R}$ is the rotation of axes matrix needed to make the eigenvector $\mathbf{c}$ the x

axis; hence $\mathbf{R} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ d_1 & d_2 \end{pmatrix}$.

$$\begin{aligned} \text{So } \mathbf{R}\mathbf{A}\mathbf{R}' &= \begin{pmatrix} c_1 & c_2 \\ d_1 & d_2 \end{pmatrix} \mathbf{A} \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix} \\ &= \begin{pmatrix} c_1 & c_2 \\ d_1 & d_2 \end{pmatrix} (\mathbf{A}\mathbf{c} \quad \mathbf{A}\mathbf{d}) \\ &= \begin{pmatrix} c_1 & c_2 \\ d_1 & d_2 \end{pmatrix} (\lambda_1 \mathbf{c} \quad \lambda_2 \mathbf{d}) \text{ since } \mathbf{A}\mathbf{c} = \lambda_1 \mathbf{c} \text{ and } \mathbf{A}\mathbf{d} = \lambda_2 \mathbf{d}, \mathbf{c} \text{ and } \mathbf{d} \\ &\quad \text{being eigenvectors,} \\ &= \begin{pmatrix} c_1 & c_2 \\ d_1 & d_2 \end{pmatrix} \begin{pmatrix} \lambda_1 c_1 & \lambda_2 d_1 \\ \lambda_1 c_2 & \lambda_2 d_2 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 (c_1^2 + c_2^2) & \lambda_2 (c_1 d_1 + c_2 d_2) \\ \lambda_1 (c_1 d_1 + c_2 d_2) & \lambda_2 (d_1^2 + d_2^2) \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \text{ using (1) and (2).} \end{aligned}$$

9.5 Eigenvectors and Principal Axes

We have seen that when the axes of coordinates are rotated into the positions of the eigenvectors of the matrix $\mathbf{A}$, the equation $\mathbf{x}'\mathbf{A}\mathbf{x} = k$ reduces to the form $\lambda_1 x_1^2 + \lambda_2 y_1^2 = k$ which is the equation of a conic (possibly degenerate) with its principal axes along the axes of coordinates. Thus we may say that the directions of the eigenvectors of $\mathbf{A}$ are the directions of the principal axes of the conic $\mathbf{x}'\mathbf{A}\mathbf{x} = k$.

9.6 The general quadratic form – translation of axes

We have, at present, only considered quadratic forms containing no first degree terms. The most general form is given by

$$f(x, y) \equiv ax^2 + 2hxy + by^2 + 2gx + 2fy + c \quad . \quad . \quad . \quad (3).$$

We have seen how the part $ax^2 + 2hxy + by^2$ can be reduced to diagonal form by a rotation of axes. What effect have the terms $2gx$ and $2fy$ on our previous analysis?

Consider the form

$$x^2 + 4xy - 2y^2 - 10x + 4y + 20 = 0 \quad (4).$$

We can use matrices to describe this form also. Check that

$$(x \ y \ 1) \begin{pmatrix} 1 & 2 & -5 \\ 2 & -2 & 2 \\ -5 & 2 & 20 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

is the same as equation (4).

Note that the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 & -5 \\ 2 & -2 & 2 \\ -5 & 2 & 20 \end{pmatrix}$ is again chosen to be symmetric.

Show that the general quadratic form (3) can be written as

$$(x \ y \ 1) \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

Make up other forms of the matrix $\mathbf{A}$ which will be equivalent to (4).

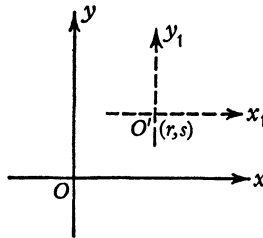


Fig. 9.4

We shall see that the terms $2gx$, $2fy$ can be removed by a *translation* of the axes, and that the resulting expression can then be reduced to diagonal form as in the preceding section.

The origin is to be translated from O to $O'(r, s)$. If the new axes are x_1, y_1 , we have

$$x_1 = x - r$$

$$y_1 = y - s.$$

These equations in matrix form are

$$\begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -r \\ 0 & 1 & -s \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

or $\mathbf{x}_1 = \mathbf{T}\mathbf{x}$ where

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & -r \\ 0 & 1 & -s \\ 0 & 0 & 1 \end{pmatrix}.$$

If we use this translation what will be the new form of the expression $\mathbf{x}'\mathbf{A}\mathbf{x}$?

By using $\mathbf{x} = \mathbf{T}^{-1}\mathbf{x}_1$, $\mathbf{x}'\mathbf{A}\mathbf{x}$ becomes

$$(\mathbf{T}^{-1}\mathbf{x}_1)'\mathbf{A}(\mathbf{T}^{-1}\mathbf{x}) = \mathbf{x}_1'[(\mathbf{T}^{-1})'\mathbf{A}\mathbf{T}^{-1}]\mathbf{x}_1 = \mathbf{x}_1'\mathbf{B}\mathbf{x}_1 \text{ say.}$$

$$\mathbf{T}^{-1} = \begin{pmatrix} 1 & 0 & r \\ 0 & 1 & s \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad (\mathbf{T}^{-1})' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ r & s & 1 \end{pmatrix}.$$

For the matrix $\mathbf{A} = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}$ $(\mathbf{T}^{-1})'\mathbf{A}\mathbf{T}^{-1}$ becomes

$$\begin{aligned} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ r & s & 1 \end{pmatrix} \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} \begin{pmatrix} 1 & 0 & r \\ 0 & 1 & s \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} a & h & ar+hs+g \\ h & b & hr+bs+f \\ ar+hs+g & hr+bs+f & ar^2+2hrs+bs^2+2gr+2fs+c \end{pmatrix} \\ &= \begin{pmatrix} a & h & ar+hs+g \\ h & b & hr+bs+f \\ ar+hs+g & hr+bs+f & f(r,s) \end{pmatrix} = \mathbf{B}. \end{aligned}$$

The symmetric form of this matrix is a consequence of taking $\mathbf{A}$ to be symmetric.

We wish to choose r and s so that the form of $\mathbf{B}$ becomes

$$px_1^2 + 2kx_1y_1 + qy_1^2 + c_1.$$

This expression in matrix form is

$$\begin{pmatrix} x_1 & y_1 & 1 \end{pmatrix} \begin{pmatrix} p & k & 0 \\ k & q & 0 \\ 0 & 0 & c_1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}.$$

Hence we chose $p=a$, $q=b$, $k=h$, $c_1=f(r,s)$ and

$$ar+hs+g=0 \quad . \quad . \quad . \quad (5)$$

$$hr+bs+f=0 \quad . \quad . \quad . \quad (6)$$

(Note the first row of $\mathbf{A} = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}$ is $a \quad h \quad g$ and the second row

$h \quad b \quad f$.)

The need for $\mathbf{A}$ to be symmetric is that the four zeros in $\mathbf{B}$ are all produced.

The equations (5) and (6) allow us to find the values of r and s , so that the quadratic form is reduced to $ax_1^2 + 2hx_1y_1 + by_1^2 + f(r, s)$. By methods described in the first section, this form may now be reduced to $mx_2^2 + ny_2^2 + c_2$ by a suitable rotation of axes and hence the locus of $f(x, y) = 0$ determined.

9.7 Worked investigations

(1) $x^2 + 4xy - 2y^2 - 10x + 4y + 20 = 0 = f(x, y)$ can be put in matrix form as

$$\begin{pmatrix} x & y & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -5 \\ 2 & -2 & 2 \\ -5 & 2 & 20 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0.$$

Let $\mathbf{A} = \begin{pmatrix} 1 & 2 & -5 \\ 2 & -2 & 2 \\ -5 & 2 & 20 \end{pmatrix}$. Then by translating the axes so that

(r, s) becomes the new origin we must have (for the quadratic form to reduce to $ax_1^2 + 2hx_1y_1 + by_1^2 + c_1 = 0$) that

$$1r + 2s - 5 = 0$$

and

$$2r - 2s + 2 = 0$$

from which $r = 1, s = 2$.

Then $f(1, 2) = 1 + 4(2) - 2(4) - 10(1) + 4(2) + 20 = 19$. Hence by translating the axes to the point $(1, 2)$ the equation becomes

$$\begin{pmatrix} x_1 & y_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 19 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = 0$$

or

$$\begin{pmatrix} x_1 & y_1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + 19 = 0.$$

The form $\begin{pmatrix} x_1 & y_1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ can now be reduced to $mx_2^2 + ny_2^2$ by a suitable rotation of axes.

Let $\mathbf{C} = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$. The eigenvalues of $\mathbf{C}$ are given by

$$(1 - \lambda)(-2 - \lambda) - 4 = 0 \quad \text{or} \quad \lambda^2 + \lambda - 6 = 0 = (\lambda + 3)(\lambda - 2)$$

that is

$$\lambda = -3 \quad \text{or} \quad 2.$$

From $(1 - \lambda)x + 2y = 0$, when $\lambda = -3$ the eigenvector is $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ and

when $\lambda = 2$, the eigenvector is $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ (as expected, perpendicular to the

first eigenvector, as $\mathbf{C}$ is symmetric). The rotation matrix to rotate the axes through $\tan^{-1} \frac{1}{2}$ to make the eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ the axis x_2 is

$$\mathbf{R} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

Hence the matrix $\mathbf{C}$ under rotation become $\mathbf{RCR}'$, or

$$\begin{aligned} \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} &= \frac{1}{5} \begin{pmatrix} 4 & 2 \\ 3 & -6 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 10 & 0 \\ 0 & -15 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix} \end{aligned}$$

(Note the eigenvalues in this diagonal matrix and the eigenvalue 2 corresponding to the eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ taken.)

The equation $(x_1 \ y_1) \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + 19 = 0$ when referred to axes of x_2, y_2 inclined at $\tan^{-1} (\frac{1}{2})$ to the x_1, y_1 axes, becomes

$$\begin{aligned} (x_2 \ y_2) \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + 19 = 0, \\ 2x_2^2 - 3y_2^2 + 19 = 0 \end{aligned}$$

or

$$-\frac{2x_2^2}{19} + \frac{3y_2^2}{19} = 1 - \text{a hyperbola.}$$

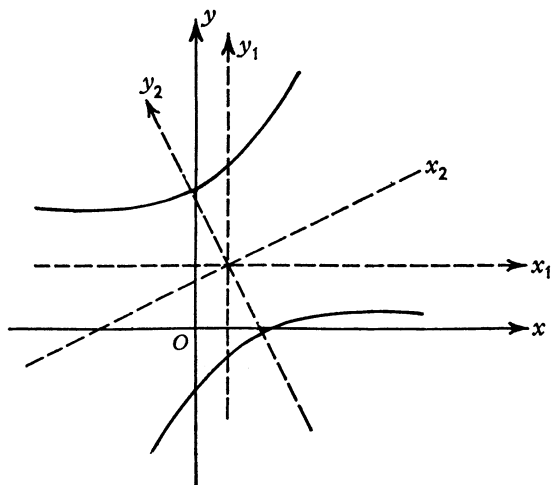


Fig. 9.5

(2) Determine the locus of the conic

$$4x^2 + 2xy + 4y^2 - 26x - 14y + 31 = 0.$$

$$f(x, y) = 4x^2 + 2xy + 4y^2 - 26x - 14y + 31 = \mathbf{x}' \begin{pmatrix} 4 & 1 & -13 \\ 1 & 4 & -7 \\ -13 & -7 & 31 \end{pmatrix} \mathbf{x} = 0.$$

$$\text{Let } \mathbf{A} = \begin{pmatrix} 4 & 1 & -13 \\ 1 & 4 & -7 \\ -13 & -7 & 31 \end{pmatrix}.$$

Consider a translation of axes so that the point (r, s) becomes the new origin. The values of r and s are given by

$$\left. \begin{array}{l} 4r + s - 13 = 0 \\ r + 4s - 7 = 0 \end{array} \right\} \text{ from which } r = 3 \text{ and } s = 1.$$

$$\text{Then } f(r, s) = f(3, 1) = -15.$$

On referring the conic to the origin $(3, 1)$ with axes x_1, y_1 , its form becomes

$$\mathbf{x}'_1 \begin{pmatrix} 4 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & -15 \end{pmatrix} \mathbf{x}_1 = 0$$

or

$$\mathbf{x}' \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix} \mathbf{x} - 15 = 0.$$

Let $\mathbf{B} = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}$. The characteristic equation of $\mathbf{B}$ is given by

$$\begin{vmatrix} 4 - \lambda & 1 \\ 1 & 4 - \lambda \end{vmatrix} = 0$$

that is

$$\begin{aligned} (4 - \lambda)^2 - 1 &= 0 \\ \lambda^2 - 8\lambda + 15 &= 0, \\ \lambda &= 3 \text{ or } 5. \end{aligned}$$

When $\lambda = 3$ the eigenvector is $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and when $\lambda = 5$ the eigenvector is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Rotating the axes to the eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, through 45° the form of the conic becomes $\mathbf{x}'_2 \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} \mathbf{x}_2 - 15 = 0$.

(Note that the eigenvalues have been put in the diagonal matrix without working out the matrices with the rotation matrix through 45° and that since the eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ was chosen corresponding to eigenvalue 5, this eigenvalue takes the top left hand corner.)

$$\mathbf{x}'_2 \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} \mathbf{x}_2 - 15 = 0$$

is

$$5x_2^2 + 3y_2^2 - 15 = 0,$$

$$\frac{x_2^2}{3} + \frac{y_2^2}{5} = 1, \text{ an ellipse.}$$

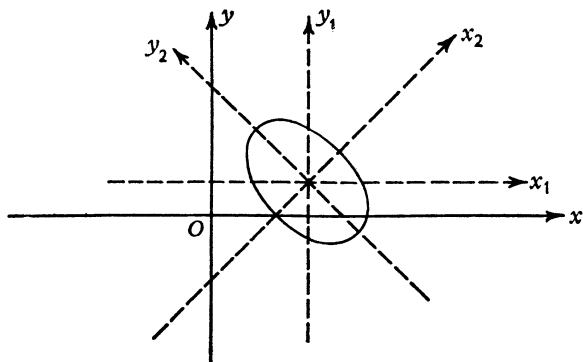


Fig. 6.6

(3) Determine the locus of the conic $x^2 + 4xy + 4y^2 - 28x - 6y - 4 = 0$.

$$f(x, y) = x^2 + 4xy + 4y^2 - 28x - 6y - 4 = \mathbf{x}' \begin{pmatrix} 1 & 2 & -14 \\ 2 & 4 & -3 \\ -14 & -3 & -4 \end{pmatrix} \mathbf{x} = 0.$$

$$\text{Let } \mathbf{A} = \begin{pmatrix} 1 & 2 & -14 \\ 2 & 4 & -3 \\ -14 & -3 & -4 \end{pmatrix}.$$

Consider a translation of axes so that the point (r, s) becomes the new origin. The values of r and s are given by

$$r + 2s - 14 = 0$$

$$2r + 4s - 3 = 0$$

which are *inconsistent* equations and hence no values of r and s .

We are thus unable to translate the axes to a suitable origin.

Let $\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. The eigenvalues are given by $(1 - \lambda)(4 - \lambda) - 4 = 0$, $\lambda^2 - 5\lambda = 0$; $\lambda = 0$ or 5 . When $\lambda = 5$ the eigenvector is $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Rotate the axes through $\tan^{-1} 2$. The matrix $\mathbf{B}$ will become $\begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}$

but the values of the other terms of $\mathbf{A}$ will also be affected by the rotation. In this particular case when there are no (r, s) values, we need to see more closely the effect of the rotation. The rotation matrix is

$\frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$ from axes xy to axes x_1y_1 say; so

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and hence

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

so that

$$x = \frac{1}{\sqrt{5}} (x_1 - 2y_1)$$

and

$$y = \frac{1}{\sqrt{5}} (2x_1 + y_1).$$

In $f(x, y)$ the terms $-28x - 6y$ become

$$\begin{aligned} & -28 \left(\frac{1}{\sqrt{5}} \right) (x_1 - 2y_1) - 6 \left(\frac{1}{\sqrt{5}} \right) (2x_1 + y_1) \\ &= \frac{1}{\sqrt{5}} (-28x_1 + 56y_1 - 12x_1 - 6y_1) \\ &= \frac{1}{\sqrt{5}} (-40x_1 + 50y_1) \\ &= -8\sqrt{5}x_1 + 10\sqrt{5}y_1. \end{aligned}$$

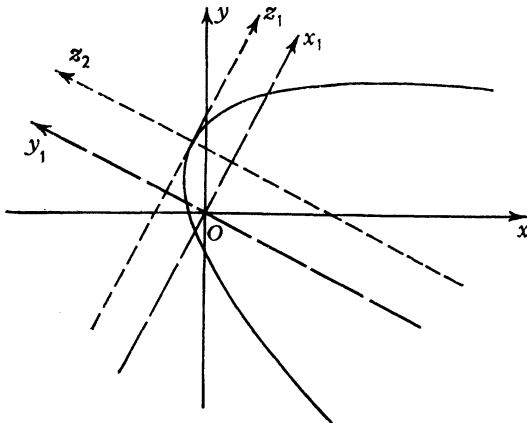


Fig. 9.7

The conic $f(x, y) = 0$ then becomes, on rotation of axes through $\tan^{-1}(2)$;

$$\mathbf{x}_1 \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{x}_1 - 8\sqrt{5}x_1 + 10\sqrt{5}y_1 - 4 = 0$$

or

$$5x_1^2 - 8\sqrt{5}x_1 + 10\sqrt{5}y_1 - 4 = 0$$

– which is a parabola, since the expression contains no y_1^2 term.

$$x_1^2 - \frac{8}{\sqrt{5}}x_1 = \frac{4}{5} - 2\sqrt{5}y_1,$$

$$\left(x_1 - \frac{4}{\sqrt{5}}\right)^2 = \frac{4}{5} - 2\sqrt{5}y_1 + \frac{16}{5} = 4 - 2\sqrt{5}y_1.$$

Let $z_1 = x_1 - \frac{4}{\sqrt{5}}$ and $z_2 = 2\sqrt{5}y_1 - 4$ then $z_1^2 = -z_2$.

9.8 The geometrical significance of the invariants in the matrices used in the reduction of quadratic forms.

If $\mathbf{x}'\mathbf{A}\mathbf{x}$, where $\mathbf{A} = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$, is a reduced quadratic form, then the geometrical characteristic of the locus $\mathbf{x}'\mathbf{A}\mathbf{x} + c = 0$ are given by the following invariants of the matrix. Let the eigenvalues of $\mathbf{A}$ be λ_1 and λ_2 then, for λ_1 and λ_2 ,

$$\begin{pmatrix} a - \lambda & h \\ h & b - \lambda \end{pmatrix} = 0$$

from which

$$(a - \lambda)(b - \lambda) - h^2 = 0,$$

$$\lambda^2 - (a + b)\lambda + ab - h^2 = 0.$$

Hence

$$\lambda_1 + \lambda_2 = a + b$$

and

$$\lambda_1\lambda_2 = ab - h^2 = \det \mathbf{A}.$$

On using a rotation matrix to make the eigenvectors the axes the diagonal form of $\mathbf{A}$ is $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, that is the quadratic form reduces to

$$\mathbf{x}_1' \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \mathbf{x}_1 + c = 0 \text{ or}$$

$$\lambda_1 x_1^2 + \lambda_2 y_1^2 + c = 0 \quad . \quad . \quad . \quad (7).$$

Now if in (7) $c=0$, the quadratic form is a line pair going through the origin of the axes x_1y_1 . If $\lambda_1 + \lambda_2 = 0$ the lines are the orthogonal pair $y_1 = \pm x_1$.

In other cases from the form of (7) we have

- (i) an ellipse if λ_1 and λ_2 have both the same sign, that is $\lambda_1\lambda_2 > 0$, that is $ab - h^2 > 0$
- (ii) a hyperbola if λ_1 and λ_2 are of different signs, that is $\lambda_1\lambda_2 < 0$, that is $ab - h^2 < 0$ (rectangular when $\lambda_1 + \lambda_2 = 0$)
- (iii) a parabola if one of λ_1 or $\lambda_2 = 0$, that is $ab - h^2 = 0$
- (iv) a circle if $\lambda_1 = \lambda_2$.

(Note that it is possible for (7) to give imaginary conics.)

Examples

Determine the locus of the following conics and sketch them

9.8 $5x^2 + 4xy + 5y^2 + 6x - 6y - 15 = 0$.

9.9 $4x^2 + 4y^2 - 16x - 7y + 4 = 0$.

9.10 $2x^2 + 4xy - y^2 + 4x - 8y - 16 = 0$.

9.11 $x^2 - 2xy + y^2 - 2x - 2y = 0$.

BIBLIOGRAPHY

School texts of 'A' level standard

Vectors, Matrices and Linear Equations, Neill, H. and Moakes, A. J. (Oliver and Boyd, 1967).

Theory and Problems of Matrices, Ayres, F. (Schaum, New York, 1962).

Academic texts

The Algebra of Vectors and Matrices, Wade, T. (Addison-Wesley, 1951).

Applications

Mechanical Vibrations: An Introduction to Matrix Methods, Prentis, J. and Leckie, F. (Longmans, 1963).

Finite Mathematics with Business Applications, Kemeny, J., Schleifer, A., Snell, J. and Thompson, G. (Prentice Hall, 1962).

Finite Graphs and Networks: An Introduction with Application, Busacker, R., and Saaty, T. (McGraw Hill, 1965).

Matrix Methods in Optical Instrument Design, Brouwer, W. (Benjamin, 1964).

Matrix Theory for Electrical Engineering Students, Tropper, A. (Harrap, 1962).

ANSWERS TO EXERCISES

Chapter 1

$$1.1 \quad \begin{pmatrix} 95 \\ 75 \\ 170 \end{pmatrix} \begin{pmatrix} -5 \\ 35 \\ 30 \end{pmatrix} \begin{pmatrix} 65 \\ 40 \\ 105 \end{pmatrix} (145, 120, 50, 315), (-10, 30, -5, 15).$$

$$1.2 \quad \begin{array}{ccccc} & \mathbf{e} & \mathbf{a} & \mathbf{f} & \mathbf{r} \\ \mathbf{S}_2 = & \mathbf{m} & 25 & 30 & 0 & 55 \\ & \mathbf{w} & 35 & 18 & 0 & 53 \\ & \mathbf{c} & 60 & 48 & 0 & 108 \end{array} \quad \begin{array}{ccccc} & \mathbf{e} & \mathbf{a} & \mathbf{b} & \mathbf{r}_3 \\ \mathbf{S}_3 = & \mathbf{m}_3 & 25 & 30 & 1 & 54 \\ & \mathbf{w}_3 & 35 & 18 & 2 & 51 \\ & \mathbf{c}_3 & 60 & 48 & 3 & 105 \end{array}$$

$$1.3 \quad \begin{array}{ccccc} & \mathbf{e} & \mathbf{a} & \mathbf{f} & \\ \mathbf{F} - \mathbf{S} = & \mathbf{m} & 20 & 20 & 15 \\ & \mathbf{w} & 20 & 2 & 20 \end{array} \quad \begin{array}{ccccc} & \mathbf{e} & \mathbf{a} & \mathbf{b} & \mathbf{r}_3 \\ \mathbf{F}_3 - \mathbf{S}_3 = & \mathbf{m}_3 & 20 & 20 & 1 & 39 \\ & \mathbf{w}_3 & 20 & 2 & 1 & 21 \\ & \mathbf{c}_3 & 40 & 22 & 2 & 60 \end{array}$$

$$4\mathbf{S} = \begin{array}{ccccc} & \mathbf{e} & \mathbf{a} & \mathbf{f} & \\ \mathbf{m} & 100 & 120 & 0 & \\ \mathbf{w} & 140 & 72 & 0. & \end{array}$$

- 1.4 English students' subscription = $45.3 + 55.2 = 245 = (45, 55) \cdot (3, 2)$
 Maths students' subscription = $50.3 + 20.2 = 190 = (50, 20) \cdot (3, 2)$
 All students' subscription = $93.3 + 72.2 = 423$
 (this is the true total).

1.5 $-6, 0, 1$

$$1.6 \quad \mathbf{AB} = \begin{pmatrix} 3 & 13 \\ 1 & -1 \end{pmatrix}, \mathbf{BA} = \begin{pmatrix} 4 & 4 \\ 2 & -2 \end{pmatrix} \text{ (note that } \mathbf{AB} \neq \mathbf{BA} \text{), } \mathbf{AC} = \mathbf{CA} = \mathbf{A},$$

$$\mathbf{AD} = \begin{pmatrix} 5 & 3 \\ -1 & 1 \end{pmatrix}, \mathbf{DA} = \begin{pmatrix} 1 & -1 \\ 3 & 5 \end{pmatrix}, \mathbf{BC} = \mathbf{CB} = \mathbf{B}, \mathbf{BD} = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix},$$

$$\mathbf{DB} = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}, \mathbf{CD} = \mathbf{DC} = \mathbf{D}; \mathbf{A}^2 = \begin{pmatrix} 14 & 10 \\ 2 & 6 \end{pmatrix}, \mathbf{B}^2 = \begin{pmatrix} 1 & 3 \\ 0 & 4 \end{pmatrix},$$

$$\mathbf{C}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

1.7 $\mathbf{EF}$ does not exist (for matrix multiplication to be possible the number of *columns* of the first matrix must be equal to the number of *rows* in the second matrix).

$$\mathbf{FE} = \begin{pmatrix} 2 & 4 & -3 \\ 9 & 6 & 15 \end{pmatrix}, \mathbf{EG} = \mathbf{GE} = \mathbf{E}, \mathbf{EH} = \begin{pmatrix} 3 & 2 & 5 \\ 1 & 1 & -1 \\ -2 & 0 & 1 \end{pmatrix}$$

$$\mathbf{HE} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 3 & 5 \\ 0 & -2 & 1 \end{pmatrix}, \mathbf{FG} = \mathbf{F}, \mathbf{GF} \text{ does not exist,}$$

$$\mathbf{FH} = \begin{pmatrix} 2 & 0 & -1 \\ 3 & 3 & 3 \end{pmatrix}, \mathbf{HF} \text{ does not exist; } \mathbf{GH} = \mathbf{HG} = \mathbf{H}.$$

$$\mathbf{E}^2 = \begin{pmatrix} 7 & -1 & 12 \\ 3 & 6 & 3 \\ -2 & -4 & 3 \end{pmatrix}. \quad \mathbf{F}^2 \text{ does not exist.}$$

$\mathbf{C}$ and $\mathbf{G}$ leave matrices unchanged; these are called unit matrices.

$\mathbf{D}$ and $\mathbf{H}$ when pre-multiplying a matrix interchange its first and second rows, when post-multiplying a matrix interchange its first and second columns.

$$\mathbf{1.8} \quad \mathbf{t} = \mathbf{Mv} \text{ where } \mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \text{ gives the } \mathbf{t} \text{ column } \begin{pmatrix} 77 \\ 131 \\ 128 \\ 207 \\ 230 \end{pmatrix}; \quad \mathbf{t}_c \mathbf{v} = 713;$$

$\mathbf{t}_s \mathbf{v} = 924$ and as $1 \cdot 30 = \frac{9 \cdot 24}{7 \cdot 13}$ the value $1 \cdot 30$ comes from using $\mathbf{mv} \mathbf{t}_c = \mathbf{t}_s$.

$$\mathbf{1.9} \quad \text{Let } \mathbf{N} \text{ be the } 6 \times 5 \text{ matrix, } \mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{u} = (1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1),$$

$$\mathbf{p} = (1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6).$$

$\mathbf{Nv}$ gives the number of families with 1 child, 2 children; the total column.

$\mathbf{uN}$ gives the number of families in each group.

$\mathbf{pN}$ gives the number of children in each group.

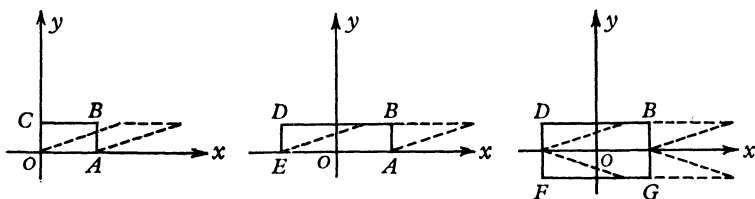
If $\mathbf{m}$ represents the row giving the mean number of children in each group, then $\mathbf{mv}(\mathbf{uN}) = (\mathbf{pN})$.

Chapter 2

2.1 (a) $\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ (c) $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

- 2.2 (a) reflection in x axis
 (b) linear stretch, factor 4 in x axis direction
 (c) enlargement, factor 3
 (d) rotation of the plane through 45°
 (e) rotation of the axes through 45° .

2.3

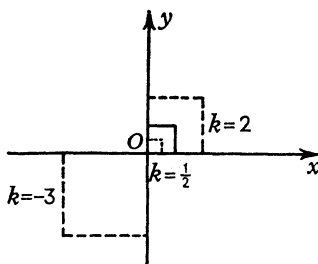


2.4 $\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$.

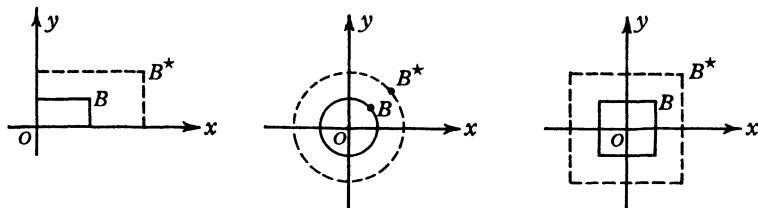
- 2.5 (i) A shear parallel to the x axis in the negative direction, factor 1.
 (ii) A shear parallel to the y axis of factor $\frac{1}{2}$.

2.6 (a) $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} = k \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

(b)

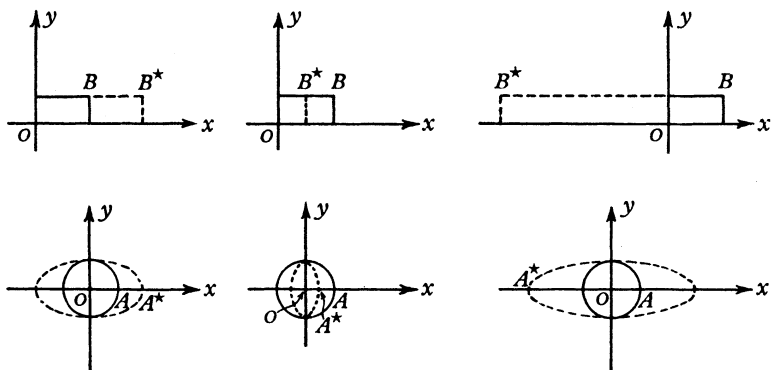


(c)



2.7 Identity, reflection in y axis, reflection in x axis, rotation through 180° .

Reflection in $x=y$, rotation through 90° , rotation through 270° , reflection in $x=-y$.

2.8

The matrix $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ transforms the plane such that $P(x, y)$ is mapped onto $P^*(x^*, y^*)$ where $\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. Then $x^* = 2x$ and $y^* = y$. Hence if (x, y) are related by the equation $x^2 + y^2 = 1$, then (x^*, y^*) are related by $\left(\frac{x^*}{2}\right)^2 + (y^*)^2 = 1$, which is an ellipse.

2.9 (a) Shear factor 2 parallel to y axis.

(b) Projection parallel to x axis onto line $y = 3x$.

2.10 $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

2.11 **M** – reflection in y axis; **Q** – rotation through 90° .

$$\mathbf{QM} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \text{ – reflection in } x = -y.$$

$$\mathbf{MQ} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ – reflection in } x = y.$$

2.12 $\mathbf{R} = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$, find the new positions of the vertices of the

unit square and plot them.

$\mathbf{S} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, use the coordinates of the rotated square and find their new positions due to the shear.

$$\mathbf{SR} = \begin{pmatrix} \frac{1}{2} & \frac{3}{2} & \frac{-\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{3\sqrt{3}}{2} + \frac{1}{2} \end{pmatrix}, \text{ find the new positions of the vertices of the}$$

original unit square and plot them. The final figures will be the same.

$$(b) \mathbf{FR} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

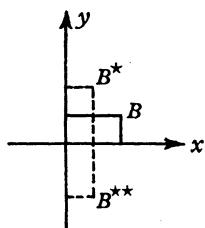
$$\mathbf{RF} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}$$

Results different.

$$(c) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 6 \\ 2 & 0 \end{pmatrix}.$$

2.13

(a)



rotation through 270°
(note position of vertices)

$$(b) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

$$2.14 \quad R_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad R_\beta = \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix}$$

$$R_{\alpha+\beta} = \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}$$

$$= R_\beta R_\alpha = \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

$$2.15 \quad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$AB = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad BA = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{results the same: trans-} \\ \text{formation of unit square} \\ \text{would correspond.}$$

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad N = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$MN = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad NM = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{results different, trans-} \\ \text{formation of unit square would} \\ \text{not correspond.}$$

$AB=MN$: a reflection in two mirrors is equivalent to a rotation through twice the angle between the mirrors. (MN implies reflection in y axis first followed by reflection in $x=y$.)

$$2.16 \quad AB=BA \text{ and } A(BC)=(AB)C.$$

$$2.17 \quad (a) \begin{pmatrix} 2 & 4 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \text{a shear followed by an en-} \\ \text{largement (or vice versa)}$$

$$(b) \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \text{a shear followed by a reflec-} \\ \text{tion in } x=y \text{ (not reversible)}$$

$$\begin{pmatrix} 0 & -3 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad \text{a double stretch followed by} \\ \text{a rotation through } 90^\circ \\ \text{(reversible)}$$

$$\begin{pmatrix} -1 & 0 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \quad \text{a shear followed by reflection} \\ \text{in } y \text{ axis (not reversible).}$$

$$2.18 \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

AB = rotation through 90°

BA = rotation through 270°

$$AB^2 = A$$

$(BA)^2$ = rotation through 180°

ABA = reflection in y axis

BAB = reflection in $x = -y$

$$ABABA = A(BA)^2 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = BAB$$

$$BABABA = B(ABABA) = B(BAB) = AB.$$

$$2.19 \quad \begin{array}{c|cc} & I & R \\ \hline I & I & R \\ R & R & I \end{array} \quad + \quad \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

$$2.20 \quad V = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad R = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

	I	R	V	H	$\times$	1	3	5	7	
I	I	R	V	H	1	1	3	5	7	
R	R	I	H	V	3	3	1	7	5	
V	V	H	I	R	5	5	7	1	3	(Mod 8)
H	H	V	R	I	7	7	5	3	1	

Chapter 3

$$3.1 \quad (i) \begin{pmatrix} 2 & -1 \\ -7 & 4 \end{pmatrix} \quad (ii) \begin{pmatrix} -2 & -1 \\ 7 & -3 \end{pmatrix} \quad (iii) \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}.$$

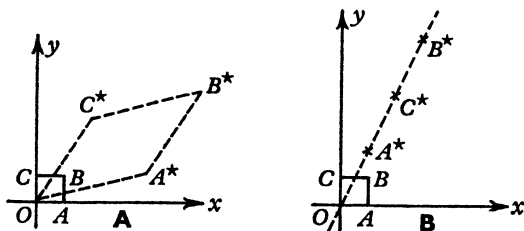
Interchange the entries in the major diagonal and reverse the sign of the entries in the minor diagonal.

$$3.2 \quad (i) -1, \begin{pmatrix} -1 & 1 \\ 3 & -2 \end{pmatrix} \quad (ii) -2, \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

$$(iii) -5, \frac{1}{5} \begin{pmatrix} -3 & 2 \\ 4 & 1 \end{pmatrix} \quad (iv) 13, \frac{1}{13} \begin{pmatrix} 5 & -2 \\ -1 & 3 \end{pmatrix}$$

$$(v) \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

3.3



The matrix $\mathbf{A}$ is a non-singular matrix which has the geometric property of 1 to 1 correspondence between points and their images. The matrix $\mathbf{B}$ is singular and gives a many-to-1 correspondence between points and their images. For $\mathbf{B}$ any point on the line $x+2y=3$ is mapped onto $(3, 6)$.

$$3.5 \quad \begin{pmatrix} \cos \alpha & -\sin \alpha & h \cos \alpha & -k \sin \alpha \\ \sin \alpha & \cos \alpha & h \sin \alpha & +k \sin \alpha \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$3.6 \quad \begin{aligned} & \text{(i) } (a+4, -b) \quad \text{(ii) } (a, -b-4) \quad \text{(iii) } (-b+4, a) \\ & \text{(iv) } (+b, +a). \end{aligned}$$

$$3.7 \quad \text{The matrix to reflect in } y=mx \text{ is } \frac{1}{1+m^2} \begin{pmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{pmatrix}.$$

The matrix to reflect in $y=mx+c$ is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} & 0 \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix} \\ = \begin{pmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} & \frac{-2mc}{1+m^2} \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} & \frac{2c}{1+m^2} \\ 0 & 0 & 1 \end{pmatrix}.$$

$$3.8 \quad \begin{aligned} & \text{(i) } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{(ii) } \begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix} \quad \text{(iii) } \begin{pmatrix} 1 & a & 0 \\ b & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

- 3.9** (i) reflection in the $x-z$ axes plane
 (ii) a linear stretch of factor a in the x axis direction
 (iii) identity
 (iv) shear factor 1 parallel to z axis.

Chapter 4

- 4.1** (a) $x=2, y=3$; (b) $x=3, y=2$; (c) $x=-\frac{1}{2}, y=1\frac{2}{3}$.
- 4.2** (i) $x=2, y=1$; (ii) $x=2, y=-1$; (iii) $x=1, y=\frac{1}{2}$;
 (iv) $x=0=y$; (v) $x=1, y=3$.
- 4.4** (i) $\begin{pmatrix} 4 & 3 & -3 \\ -4 & -2 & 3 \\ -1 & -1 & 1 \end{pmatrix}$ (ii) $\frac{1}{4} \begin{pmatrix} -5 & 7 & 1 \\ -1 & -1 & 1 \\ -3 & -5 & 1 \end{pmatrix}$
 (iii) $\frac{1}{6} \begin{pmatrix} -5 & 9 & -3 \\ 3 & -3 & -3 \\ -7 & 9 & -3 \end{pmatrix}$ (iv) $\begin{pmatrix} 5 & -7 & -2 & -1 \\ 7 & -10 & -3 & -2 \\ -15 & 22 & 6 & 4 \\ -3 & 4 & 1 & 1 \end{pmatrix}$.

Chapter 5

- 5.3** (a) (i) 1, (ii) 2, (iii) 0, (iv) -3, (v) 1.
 (b) (i) and (v).
- 5.4** (i) 1, 6; $x-4y=0, x+y=0$; $\begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.
 (ii) 5, 10; $x+2y=0, 2x-y=0$; $\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.
- 5.5** 2, 1; $\pm i$.
- 5.6** $\lambda^3 - \lambda(a+d) + ad - bc = 0$,
 Sum of eigenvalues is sum of major diagonal entries, product of eigenvalues is determinant value.
- 5.7** See 5.5(ii).
- 5.8** $\begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$ has repeated eigenvalues of 2.

5.9 Eigenvalues are major diagonal entries.

5.10 k and k .

5.11 $a \pm b$.

5.12 $7, 2; 7, 2; -5, 8; -5, 8$.

5.13 $x + 3y = 0, x - y = 0; x$ axis, $x = y; x + 2y = 0$, complex.

- 5.14** (a) real distinct eigenvectors
 (b) complex eigenvectors
 (c) any line through the origin is an eigenvector
 (d) one eigenvector
 (e) x and y axes as eigenvectors.

5.15 (i) $1, 2, 3; \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$.

(ii) $-1, 2, 1; \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$.

5.16 $\mathbf{AB} = \begin{pmatrix} 4 & 17 \\ 1 & 38 \end{pmatrix} \quad \mathbf{A'B'} = \begin{pmatrix} 18 & 27 \\ 11 & 24 \end{pmatrix} \quad \mathbf{B'A'} = \begin{pmatrix} 4 & 1 \\ 17 & 38 \end{pmatrix}$

$(\mathbf{AB})' = \begin{pmatrix} 4 & 1 \\ 17 & 38 \end{pmatrix}$. Hence $(\mathbf{AB})' = \mathbf{B'A'}$.

5.17 Eigenvalues 3 or -7 , eigenvector lines $x - 2y = 0, 2x + b = 0$. The eigenvectors are orthogonal.

The characteristic equation for $\mathbf{M} = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ is

$$\lambda^2 - \lambda(a + b) + ab - c^2 = 0.$$

For the equation to have real *and* distinct eigenvalues

$$(a + b)^2 - 4(ab - c^2) > 0.$$

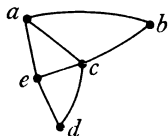
The left hand side is $(a - b)^2 + 4c^2$ which is greater than zero for all a, b, c . If $\mathbf{u}$ and $\mathbf{v}$ are the eigenvectors with corresponding eigenvalues λ_1 and λ_2 , then $\mathbf{Mu} = \lambda_1\mathbf{u}$ and $\mathbf{Mv} = \lambda_2\mathbf{v}$.

From $\mathbf{Mu} = \lambda_1\mathbf{u}$, then $(\mathbf{Mu})' = \lambda_1\mathbf{u}'$ or $\mathbf{u'M'} = \lambda_1\mathbf{u}'$, since $(\mathbf{Mu})' = \mathbf{u'M'}$ (see 5.16). As $\mathbf{M} = \mathbf{M'}$, $\mathbf{u'M} = \lambda_1\mathbf{u}'$. Postmultiplying by $\mathbf{v}$, $\mathbf{u'Mv} = \lambda_1\mathbf{u'v}$ and as $\mathbf{Mv} = \lambda_2\mathbf{v}$, $\mathbf{u'\lambda_2v} = \lambda_1\mathbf{u'v}$. Thus $\lambda_2\mathbf{u'v} - \lambda_1\mathbf{u'v} = 0$ or $(\lambda_2 - \lambda_1)\mathbf{u'v} = 0$ and since $\lambda_1 \neq \lambda_2$ (eigenvalues are real and distinct), $\mathbf{u'v} = 0$, showing the eigenvectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

$$5.18 \quad \mathbf{M}^{-1} = \begin{pmatrix} 4 & 3 & -3 \\ -4 & -2 & 3 \\ -1 & -1 & 1 \end{pmatrix}.$$

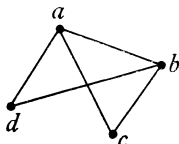
Chapter 7

7.1



$$\begin{array}{c} a \quad b \quad c \quad d \quad e \\ a \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix} \\ b \\ c \\ d \\ e \end{array}$$

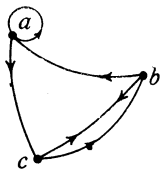
7.2



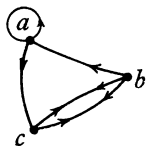
The matrices are symmetrical since if a is joined to b , b is joined to a . The sum of the individual rows (or columns) give the number of connections to (or from) a node.

When the matrix is not symmetrical the sum of entries in a row give the number of connections *from* that node to other nodes and the sum of the entries in the columns give the number of connections *to* that node.

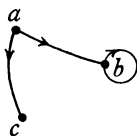
7.3



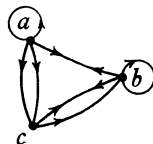
$$7.4 \quad \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & 2 & 0 \end{pmatrix}.$$



with



gives



- 7.5 (a) the number of links to a station
 (b) each link connects two stations.

7.6 Rows and columns containing two 1's—rows b, c and columns 3, 4; also rows a, b, c and columns 1, 2, 3.

- 7.7 (a) 1 and 2 from the a row.
 (b) Links 2, 3 and 4 cut, but not 1. The links to cut are found by adding modulo 2 (so $1 + 1 = 0$), the entries in the columns of the rows a and b , and cutting that link when the result is 1.

$$7.8 \quad \mathbf{A}' = \begin{matrix} & a & b & c & d \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix} \quad \mathbf{AA}' = \begin{matrix} & a & b & c & d \\ \begin{matrix} a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 3 & 2 & 0 \\ 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

The entries in $\mathbf{AA}'$ give the number of links between different stations; the leading diagonal giving the number of links at that station.

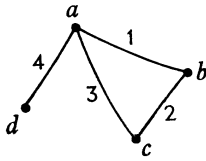
$$7.9 \quad \mathbf{P} = \begin{matrix} & \alpha & \beta & \gamma \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \mathbf{Q} = \begin{matrix} & \alpha & \beta & \gamma \\ \begin{matrix} a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

$$\mathbf{PP}' = \begin{matrix} & 1 & 2 & 3 & 4 & 5 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 2 & 2 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{pmatrix} \end{matrix} \quad \begin{array}{l} \text{The entries are the number of} \\ \text{regions common to the links; the} \\ \text{leading diagonal giving the number} \\ \text{of regions with each link.} \end{array}$$

$$\mathbf{P}'\mathbf{P} = \begin{matrix} & \alpha & \beta & \gamma \\ \begin{matrix} \alpha \\ \beta \\ \gamma \end{matrix} & \begin{pmatrix} 3 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 4 \end{pmatrix} \end{matrix} \quad \begin{array}{l} \text{The entries are the number of links com-} \\ \text{mon to the regions; the leading diagonal} \\ \text{giving the number of links with each} \\ \text{region.} \end{array}$$

$$\mathbf{QQ}' = \begin{matrix} & a & b & c & d \\ \begin{matrix} a \\ b \\ c \\ d \end{matrix} & \begin{pmatrix} 2 & 2 & 2 & 1 \\ 2 & 3 & 3 & 1 \\ 2 & 3 & 3 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \end{matrix} \quad \begin{array}{l} \text{The entries give the number of} \\ \text{regions common to the stations; the} \\ \text{leading diagonal giving the number of} \\ \text{regions with each station.} \end{array}$$

7.10 (i)



(ii) not possible, since the column of each link must have two, and only two 1's indicating the two connected stations.

7.11 $M =$

$$\begin{matrix} & a & b & c & d & e \\ \begin{matrix} a \\ b \\ c \\ d \\ e \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

 $M^2 =$

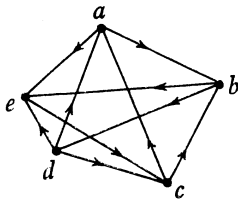
$$\begin{matrix} & a & b & c & d & e \\ \begin{matrix} a \\ b \\ c \\ d \\ e \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \end{pmatrix} \end{matrix}$$

Powers ($M + M^2$)

$$\begin{matrix} a & 2 \\ b & 4 \\ c & 2 \\ d & 3 \\ e & 5 \end{matrix}$$

e has the most 1 and 2 stage connections with the other stations.

7.12 (i)

(ii) $A^2 =$

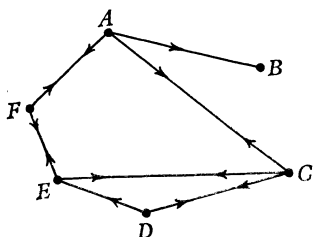
$$\begin{matrix} & a & b & c & d & e \\ \begin{matrix} a \\ b \\ c \\ d \\ e \end{matrix} & \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ 1 & 2 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

(iii) Yes. (iv) $A + A^2 =$

$$\begin{matrix} & a & b & c & d & e & \text{Powers are} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \end{matrix} & \begin{pmatrix} 0 & 1 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 & 2 \\ 1 & 2 & 0 & 1 & 2 \\ 2 & 2 & 2 & 0 & 2 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix} & \begin{matrix} a & 5 \\ b & 6 \\ c & 6 \\ d & 8 \\ e & 3 \end{matrix} \end{matrix}$$

d can communicate with all others (twice) in one or two stages. Only e cannot communicate (with d) in one or two stages.

7.13



$$\mathbf{M} = \begin{matrix} & \begin{matrix} A & B & C & D & E & F \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \\ E \\ F \end{matrix} & \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

$$\mathbf{M}^2 = \begin{matrix} & \begin{matrix} A & B & C & D & E & F \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \\ E \\ F \end{matrix} & \begin{pmatrix} 2 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 2 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 2 & 0 & 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 & 0 & 2 \end{pmatrix} \end{matrix}$$

Powers $\mathbf{M} + \mathbf{M}^2$ are

A	8
B	0
C	10
D	7
E	7
F	7

Mrs. C will spread a rumour quickest throughout the group.

7.14

$$\mathbf{A} = \begin{matrix} & \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix} \end{matrix}$$

$$\mathbf{R} = \begin{matrix} & \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix} & \begin{pmatrix} 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

$$\mathbf{A}^2 = \begin{matrix} & \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix} & \begin{pmatrix} 2 & 2 & 3 & 2 & 0 & 2 \\ 2 & 3 & 2 & 2 & 1 & 2 \\ 2 & 2 & 3 & 2 & 0 & 2 \\ 1 & 3 & 2 & 4 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 2 & 3 & 1 & 4 \end{pmatrix} \end{matrix}$$

$$\mathbf{A} + \mathbf{A}^2 + \mathbf{R} = \begin{matrix} & \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix} & \begin{pmatrix} 2 & 2 & 2 & 3 & 0 & 3 \\ 2 & 3 & 3 & 3 & 0 & 3 \\ 2 & 3 & 3 & 3 & 1 & 3 \\ 2 & 4 & 3 & 4 & 0 & 4 \\ -1 & 0 & 0 & -1 & 0 & 0 \\ 2 & 4 & 3 & 4 & 1 & 4 \end{pmatrix} \end{matrix} \begin{matrix} 12 \\ 14 \\ 15 \\ 17 \\ -2 \\ 18 \end{matrix}$$

$$\begin{matrix} & \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \end{matrix} & \begin{pmatrix} 9 & 16 & 14 & 16 & 2 & 17 \end{pmatrix} \end{matrix}$$

$$\begin{aligned} \text{Social index } a &= 9 + 2(12) = 33 \\ b &= 16 + 2(14) = 44 \\ c &= 14 + 2(15) = 44 \\ d &= 16 + 2(17) = 50 \\ e &= 2 + 2(-2) = -2 \\ f &= 17 + 2(18) = 55. \end{aligned}$$

Chapter 8

8.1 3,000 – 1st year; 6,916 (approx.) 2nd year; = 5,666; 3,000;

8.2 (a) $\frac{2}{3}$ of the species survive their first birthday; of these $\frac{1}{4}$ survive their second birthday and live into their third year and none of the species live longer than 3 years. In the third year 6 of the species are born to each alive. $\mathbf{M}^3 = \mathbf{I}$.

(b) $\frac{1}{3}$ of the species survive their first birthday and live into their second year; none of the species live longer than 2 years. In the second year, 3 of the species are born for each alive. $\mathbf{M}^2 = \mathbf{I}$.

(c) $\frac{1}{3}$ of the species survive their first birthday, of these $\frac{1}{2}$ survive their second and of these $\frac{1}{4}$ survive their third and none of the species live longer than 4 years. In the fourth year 24 of the species are born for each alive. $\mathbf{M}^4 = \mathbf{I}$.

8.3 100 of the species alive in the first and second year of life as a starting position.

$$\begin{array}{rcl} & \text{Total} & \\ \begin{pmatrix} 0 & 6 \\ \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & = & \begin{pmatrix} 6 \\ \frac{2}{3} \end{pmatrix} \quad 6\frac{2}{3} \\ \begin{pmatrix} 0 & 6 \\ \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} 6 \\ \frac{2}{3} \end{pmatrix} & = & \begin{pmatrix} 4 \\ 4 \end{pmatrix} \quad 8 \\ \begin{pmatrix} 0 & 6 \\ \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \end{pmatrix} & = & \begin{pmatrix} 24 \\ \frac{8}{3} \end{pmatrix} \quad 26\frac{2}{3} \\ \begin{pmatrix} 0 & 6 \\ \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} 24 \\ \frac{8}{3} \end{pmatrix} & = & \begin{pmatrix} 16 \\ 16 \end{pmatrix} \quad 32. \end{array}$$

The matrix $\begin{pmatrix} 0 & 6 \\ \frac{2}{3} & 0 \end{pmatrix}$ describes a species with increasing total population.

For the matrix $\begin{pmatrix} 0 & b \\ a & 0 \end{pmatrix}$ (i) $ab = 1$ gives a population which is steady, having the same values in alternate years (ii) $ab > 1$ the population increases (iii) $ab < 1$ the population decreases.

$$8.4 \quad \mathbf{P}^{-1} = \frac{1}{c_1 d_2 - c_2 d_1} \begin{pmatrix} d_2 & -d_1 \\ -c_2 & c_1 \end{pmatrix}.$$

Now $\mathbf{TP} = \mathbf{T} \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix}$ and since $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ is an eigenvector of $\mathbf{T}$, then $\mathbf{T} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 c_1 \\ \lambda_1 c_2 \end{pmatrix}$ and similarly $\mathbf{T} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} \lambda_2 d_1 \\ \lambda_2 d_2 \end{pmatrix}$.

$$\text{Hence } \mathbf{TP} = \begin{pmatrix} \lambda_1 c_1 & \lambda_2 d_1 \\ \lambda_1 c_2 & \lambda_2 d_2 \end{pmatrix}.$$

$$\begin{aligned} \text{Then } \mathbf{P}^{-1}\mathbf{TP} &= \frac{1}{c_1 d_2 - c_2 d_1} \begin{pmatrix} d_2 & -d_1 \\ -c_2 & c_1 \end{pmatrix} \begin{pmatrix} \lambda_1 c_1 & \lambda_2 d_1 \\ \lambda_1 c_2 & \lambda_2 d_2 \end{pmatrix} \\ &= \frac{1}{c_1 d_2 - c_2 d_1} \begin{pmatrix} \lambda_1 c_1 d_2 - \lambda_1 d_1 c_2 & 0 \\ 0 & -\lambda_2 d_1 c_2 + \lambda_2 d_2 c_1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \end{aligned}$$

$$\begin{aligned} 8.5 \quad \mathbf{T}^n &= \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \frac{1}{c_1 d_2 - c_2 d_1} \begin{pmatrix} d_2 & -d_1 \\ -c_2 & c_1 \end{pmatrix} \\ &= \frac{1}{c_1 d_2 - c_2 d_1} \begin{pmatrix} c_1 d_2 \lambda_1^n - d_1 c_2 \lambda_2^n & c_1 d_1 (\lambda_1^n - \lambda_2^n) \\ c_2 d_2 (\lambda_1^n - \lambda_2^n) & -c_2 d_1 \lambda_1^n + c_1 d_2 \lambda_2^n \end{pmatrix}. \end{aligned}$$

If $\lambda_1 = 1$ and $\lambda_2 < 1$ then $\lambda_2^n \rightarrow 0$ as $n \rightarrow \infty$ and a finite limit for $\mathbf{T}^n$ exists. If $\lambda_1 > 1$ the population increases and if $\lambda_1 < 1$, $\lambda_2 < 1$ then $\lambda_1^n \rightarrow 0$, $\lambda_2^n \rightarrow 0$ as $n \rightarrow \infty$ and $\mathbf{T}^n$ is zero.

8.6 (a) Steady state; eigenvalues 1 and $-\frac{1}{2}$. If the start was 100 to each life year, the steady state is 300 and 50 in the first and second year of life.

(b) eigenvalues $\frac{3}{4}$ and $-\frac{1}{2}$. Population becomes extinct.

(c) eigenvalues $1\frac{1}{2}$ and $-\frac{1}{2}$. Population increases.

8.8
$$\begin{array}{cc} & C \quad T \text{ first year} \\ \text{second} & C \begin{pmatrix} \frac{1}{2} & \frac{2}{3} \end{pmatrix} \\ \text{year} & T \begin{pmatrix} \frac{1}{2} & \frac{1}{3} \end{pmatrix} \end{array} \quad \text{eigenvalues } 1 \text{ and } -\frac{1}{6}.$$

Steady state of 8,000 country dwellers, 6,000 town dwellers.

8.9
$$\begin{array}{cc} & L \quad E \text{ first lecture} \\ \text{second} & L \begin{pmatrix} \frac{1}{4} & \frac{1}{3} \end{pmatrix} \\ \text{lecture} & E \begin{pmatrix} \frac{3}{4} & \frac{2}{3} \end{pmatrix} \end{array} \quad \text{eigenvalues } 1, -\frac{1}{12}.$$

Probability of $\frac{4}{13}$ late for end of term lecture.

8.10 As matrices have probabilities as entries, the column considers all outcomes and then must total 1.

Matrices are of the form $\begin{pmatrix} a & 1-b \\ 1-a & b \end{pmatrix}$ and

$$\begin{pmatrix} a & 1-b \\ 1-a & b \end{pmatrix} \begin{pmatrix} a & 1-b \\ 1-a & b \end{pmatrix} = \begin{pmatrix} a^2+1-a-b+ab & a-ab+b-b^2 \\ a^2-a+b-ab & 1-a-b+ab+b^2 \end{pmatrix}$$

giving column totals of 1 again.

8.12
$$\begin{array}{cc} & L \quad E \quad OT \\ L & \begin{pmatrix} 0 & \frac{2}{3} & \frac{2}{3} \end{pmatrix} \\ E & \begin{pmatrix} 1 & 0 & \frac{1}{3} \end{pmatrix} \\ OT & \begin{pmatrix} 0 & \frac{1}{3} & 0 \end{pmatrix} \end{array} \quad \text{Probabilities for end of term lecture: } L, \frac{8}{20}; E, \frac{9}{20}; \text{ On Time } \frac{3}{20}.$$

8.13
$$\mathbf{T} = \text{to} \begin{array}{cc} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \text{ from} \\ \begin{pmatrix} 0 & \frac{3}{5} & 0 & 0 \\ 1 & 0 & \frac{2}{3} & 0 \\ 0 & \frac{2}{5} & 0 & 1 \\ 0 & 0 & \frac{1}{3} & 0 \end{pmatrix} \end{array}$$

Probabilities of being in cells: 1, $\frac{9}{54}$; 2, $\frac{25}{54}$; 3, $\frac{15}{54}$; 4, $\frac{5}{54}$.

Chapter 9

9.1 All are $6x^2 + 4xy + 3y^2 = 1$.

9.2
$$\mathbf{A} = \begin{pmatrix} 6 & 5 \\ -2 & 3 \end{pmatrix}.$$

9.3
$$\mathbf{x}' \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} \mathbf{x} = 1; \quad \mathbf{x}' \begin{pmatrix} 4 & -3 \\ 0 & 2 \end{pmatrix} \mathbf{x} = 1.$$

- 9.4 Ellipse $\frac{x_1^2}{4} + \frac{y_1^2}{2} = 1$, x_1 axis gradient 1.
- 9.5 Hyperbola $\frac{x_1^2}{3} - \frac{y_1^2}{2} = 1$, x_1 axis gradient $\frac{1}{2}$.
- 9.6 Straight lines $y_1^2 = 4$,
 x_1 axis gradient 2.
- 9.7 Circle – note many eigenvectors obtained.
- 9.8 Ellipse centre $(-1, 1)$; axes 45° , $x_2^2/3 + y_2^2/7 = 1$.
- 9.9 Circle, centre $(2, 1)$ radius 2.
- 9.10 Hyperbola; centre $(1, -2)$ axes $\tan^{-1}(\frac{1}{2})$, $x_2^2/2 - y_2^2/3 = 1$.
- 9.11 Parabola, axes 45° , $y_1^2 = \sqrt{2}x_1$.